Computer Vision and Hybrid Machine Learning Algorithms on Glioblastoma 3D Cell Culture Multiplex Environment

by

Banu Erdem

A Dissertation Submitted to the
Graduate School of Health Sciences
in Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy

in

Neuroscience

KOÇ ÜNİVERSİTESİ

25 December 2025

Computer Vision and Hybrid Machine Learning Algorithms on Glioblastoma 3D Cell Culture Multiplex Environment

Koç University

Graduate School of Health Sciences

This is to certify that I have examined this copy of a doctoral dissertation by

Banu Erdem

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

Asst. Prof. Emel Sokullu (Advisor)

Prof. Dr. İhsan Solaroğlu

Assoc. Prof. Göktuğ Akyoldaş

Prof. Dr. Gökhan Bora Esmer

Prof. Dr. Gülgün Şengül

Date: 25 December 2025

To my beloved family, whose unwavering
support and encouragement
have formed the foundation of my journey.
To my close friends, for their support, empathy,
and companionship throughout this journey.
To my advisors and teachers,
whose guidance and expertise
have profoundly shaped my academic path.
This work is dedicated, with deep gratitude,
to all those who encouraged me to pursue
knowledge and excellence.

ABSTRACT

Computer Vision and Hybrid Machine Learning Algorithms
on Glioblastoma 3D Cell Culture Multiplex Environment

Banu Erdem

Doctor of Philosophy in Neuroscience
25 December 2025

Artificial intelligence (AI) is primarily comprised of machine learning and deep learning computational methodologies, offering substantial promise for improving the therapeutic management of glioblastoma, the most prevalent and aggressive adult brain tumor in the central nervous system. The objective of this study was to predict the migration and proliferation behavior of U87 human glioblastoma cells interacting with I-NHA human astrocytes. These engineered cell models, embedded within specific hydrogels, enabled us to mimic the organisms and observe their interactions. Our research involved a significant amount of raw data obtained from classical software calculations combined with image datasets captured from three-dimensional (3D) tumor microenvironment studies. *In vitro* studies indicate that glioblastoma exhibits distinct cellular characteristics, including migration, invasion, and proliferation behaviors. In this context, this work evaluated the effectiveness of Recurrent Neural Network (RNN) algorithms on 300 normalized image data obtained from two technical replicates and calculated the corresponding parameters as AI features. Specifically, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models of RNN, along with their hybrid variants with Convolutional Neural Network (CNN), were utilized to apply machine learning techniques to cell culture images. Among these models, GRU demonstrated superior estimation performance, particularly when the data density decreased; the optimized GRU model achieved a maximum accuracy of 0.85. While transformer-based models are highly effective for processing large-scale datasets due to their capacity to encode long-range dependencies, RNN-based approaches were observed to outperform transformers in time-series analyses involving medium-sized cell image datasets. These findings emphasize that enhancing image data quality and augmenting datasets with both technical and biological replicates could enhance the performance of machine learning algorithms. In summary, while this investigation demonstrates encouraging predictive accuracy, it serves primarily as a proof of concept. AI-based analyses in glioblastoma research have demonstrated promising results; further investigation and integration of larger and more diverse data are required to standardize methodology and improve the interpretability and generalizability of the prediction results.

ÖZETÇE

Üç Boyutlu Glioblastoma Çoklu Hücre Kültürü Ortamında
Görüntü İşleme ve Hibrit Makine Öğrenmesi Algoritmaları

Banu Erdem
Sinirbilim Doktora Programı
25 Aralık 2025

Yapay zeka, esas olarak makine öğrenmesi ve derin öğrenme hesaplama yöntemlerinden oluşmakta olup, merkezi sinir sisteminde en yaygın ve yetişkinlerde en hızlı ilerleyen beyin tümörü olan glioblastomanın tedavi yönetimini iyileştirme konusunda önemli bir potansiyel sunmaktadır. Bu çalışmanın amacı, I-NHA insan astrositleri ile etkileşim halinde olan U87 insan glioblastoma hücrelerinin göç ve proliferasyon davranışlarını makine öğrenmesi yöntemleri kullanarak tahmin etmektir. Belirli hidrojeller içerisine gömülmüş bu mühendislik ürünü hücre modelleri, mimetik mikroçevrelerin oluşturulmasını ve hücreler arasındaki etkileşimlerin gözlemlenmesini mümkün kılmaktadır. Araştırmamız, klasik yazılım hesaplamalarından elde edilen ve üç boyutlu (3B) tümör mikroçevresi çalışmalarından elde edilen görüntü veri setlerini içeren önemli miktarda ham veriyi kapsamaktadır. *In vitro* araştırmalara göre, glioblastomalarda göç, invazyon ve proliferasyon gibi kendine özgü hücresel özellikler ve davranışlar gözlenmektedir. Uygulamalı fizik perspektifinden bakıldığında, bu çalışma hücrelerin fiziksel davranışlarını 3B çoklu platformlarda yeniden üreterek, yapay zeka tabanlı analizler için uygun geniş kapsamlı veri setlerinin elde edilmesini sağlamaktadır. Bu araştırmada, iki teknik deney tekrarından elde edilen, normalize edilmiş 300 görüntüden oluşan veri seti üzerinde Yinelenen Sinir Ağı (*Recurrent Neural Network*, RNN) algoritmalarının etkinliği değerlendirilmiş ve ilgili parametrelerin tahmini hesaplanmıştır. Özellikle RNN'in Uzun Kısa Süreli Bellek (*Long Short-Term Memory*, LSTM) ve Kapılı Tekrarlayan Birim (*Gated Recurrent Unit*, GRU) modelleri ile ara katmanlı bir mimari sunan Konvolüsyonel Sinir Ağı (*Convolutional Neural Networks*, CNN) ile oluşturulmuş hibrit varyantları kullanılarak, hücre kültürü görüntüleri üzerinde makine öğrenimi teknikleri uygulanarak yüksek doğrulukta tahminler elde edilmiştir. Bu modeller arasında GRU, özellikle veri yoğunluğunun azaldığı durumlarda üstün bir öngörü performansı sergilemiş olup optimize edilmiş GRU modeli en yüksek 0,85 doğruluk değerine ulaşmıştır. Daha uzun dizileri kodlayabilme yetenekleri sayesinde transformer modeli büyük veri setlerinde öne çıkarken, zaman serisi tabanlı orta boyutlu hücre görüntü veri setlerinde genellikle RNN'ler daha iyi performans göstermektedir. Çalışmamızda, görüntü verisinin iyileştirilmesi ve teknik ile biyolojik çoğaltmanın, makine öğrenmesi algoritmalarının tahmin doğruluğunu önemli ölçüde artırabileceği savunulmuştur. Sonuç olarak, bu çalışma umut verici öngörü doğruluğu sergilemekle birlikte, öncelikle bir kavram kanıtlama çalışmasıdır. Glioblastoma araştırmalarında yapay zeka odaklı analizler parlak sonuçlar sunmuş olup modellerin standartlaştırılması ve öngörü sonuçlarının yorumlanabilirliğinin ve genellenebilirliğinin artırılması için daha kapsamlı ve çeşitli veri setlerinin entegrasyonu ile ileri çalışmalara ihtiyaç bulunmaktadır.

ACKNOWLEDGMENTS

I wish to express my profound appreciation to my supervisor, Asst. Prof. Dr. Emel Sokullu, for her valuable guidance, encouragement, and unwavering support throughout my doctoral journey.

I extend my sincere gratitude to my committee members, Prof. Dr. İhsan Solaroğlu and Prof. Dr. Gökhan Bora Esmer, for their profound insights, distinguished and constructive comments, and multidisciplinary perspectives, which have considerably enriched this research within the academic milieu.

I convey heartfelt thanks to Meltem Balaban and my thesis jury members, Assoc. Prof. Göktuğ Akyoldaş and Prof. Dr. Gülgün Şengül, for their scientific and methodological guidance, whose experience and recommendations were crucial in advancing this thesis.

I am writing to express my sincere gratitude to all the professors in the Neuroscience Ph.D. program, whose mentorship, knowledge, and dedication have profoundly influenced my academic journey and deepened my passion for this field of research.

My gratitude also extends to all my lab fellows at Koç University, both in Rumelifeneri and Topkapı, as well as at Koç University Hospital. Specifically, I would like to express my sincere gratitude to all the graduate and undergraduate members of the SOLAB. Their combined support, motivation, and intellectual discourse have profoundly influenced my perspective. I am especially appreciative of Baseerat, Ferdi, Nasir, and Nilufar for their tremendous assistance and collaboration during this endeavor.

I extend my sincere gratitude to my friends and lecturers at the Koç University Student Interest Group in Neurology, whose stimulating conversations and friendly environment notably enhanced my academic experience. I am profoundly grateful to the Painting Club lecturer and members, whose creativity and support provided equilibrium, inspiration, and a significant artistic avenue during my studies.

I am especially grateful for the constant encouragement of my family and close friends, whose unwavering support has provided me with the strength and motivation I need during challenging times.

I would also like to express my heartfelt appreciation to all my colleagues who supported me throughout this journey. They consistently showed genuine encouragement and kindness, always celebrating my achievements. Their unwavering positivity and camaraderie created a genuinely inspiring and uplifting environment, for which I am sincerely grateful.

I extend my appreciation to the Koç University Research Center for Translational Medicine, KUTTAM, for providing essential resources and infrastructure, funded by the Presidency of Türkiye, Head of Strategy and Budget. The facilities and collaborative environment at KUTTAM have been crucial for the completion of this dissertation.

This achievement would not have been possible without the inspiration, encouragement, and support of all those mentioned above.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

2D	Two-Dimensional
3D	Three-Dimensional
A172	A172 *Homo sapiens* brain glioblastoma cell line
AD	Alzheimer's Disease
ADNI	Alzheimer's Disease Neuroimaging Initiative
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
ANN	Artificial Neural Network
AUC	Area Under the Curve
BBB	Blood Brain Barrier
BMEC	Brain Microvascular Endothelial Cell
BT	Brain Tumor
CCLE	Cancer Cell Line Encyclopedia
cDNA	Complementary Deoxyribonucleic Acid
CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
CHB-MIT	Children's Hospital Boston–Massachusetts Institute of Technology
CI	Confidence Interval
CNN	Convolutional Neural Network
CNS	Central Nervous System
CT	Computed Tomography
CTRL	Control (untreated group)
CTRP	Cancer Therapeutics Response Portal
DeiT	Data-Efficient Image Transformer
DL	Deep Learning
DNN	Deep Neural Network
ECM	Extracellular Matrix
EEG	Electroencephalography
FFPE	Formalin-Fixed Paraffin-Embedded

GAN	Generative Adversarial Network
GAT	Graph Attention Network
GB	Glioblastoma
GBS	Glioblastomas (when used as plural GBs)
GCN	Graph Convolutional Network
GelMA	Gelatin Methacrylate
GFAP	Glial Fibrillary Acidic Protein
GFP	Green Fluorescent Protein
GNN	Graph Neural Network
GRU	Gated Recurrent Unit
h	Hours
HAMA	Hyaluronic Acid Methacrylate
HD	Huntington's disease
HSV	Hue, Saturation, Value
ICGC	International Cancer Genome Consortium
IHC	Immunohistochemistry
IMC	Imaging Mass Cytometry
iPSC	Induced Pluripotent Stem Cell
LGG	Lower-Grade Glioma
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
mg	Milligram
µm	Micrometer
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MN	Motor Neuron
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (cell viability and proliferation assay)
NGS	Next-Generation Sequencing
NHA / l-NHA	Normal Human Astrocyte / Immortalized Normal Human Astrocyte
NMJ	Neuromuscular Junction
OTSU	Otsu's Thresholding Method

PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PD	Parkinson’s disease
PDAC	Pancreatic Ductal Adenocarcinoma
PDC	Patient-Derived Cells
PDT	Patient-Derived Tumor
PEG	Polyethylene Glycol
PI	Prediction Interval
RAPD	Relative Afferent Pupillary Defect
ReLU	Rectified Linear Unit
RFP	Red Fluorescent Protein
RGB	Red, Green, Blue
RHUH-GB	Rennes University Hospital Glioblastoma Dataset
RMSE	Root Mean Squared Error
RNA	Ribonucleic Acid
RNN	Recurrent Neural Network
RSL	RAS Synthetic Lethal
RT-qPCR	Reverse Transcription Quantitative PCR
RUL	Remaining Useful Life
SCO	Spinal Cord Organoid
SE	Standard Error
ssGSEA	Single-Sample Gene Set Enrichment Analysis
TAM	Tumor-Associated Macrophages
TCGA-GB	The Cancer Genome Atlas Glioblastoma
TCIA	The Cancer Imaging Archive
TME	Tumor Microenvironment
TMZ	Temozolomide
U87	U87 *Homo sapiens* brain glioblastoma cell line
VECNN	Vision Transformer–Integrated Convolutional Neural Network
ViT	Vision Transformer
WHO	World Health Organization

CHAPTER 1: INTRODUCTION

In mathematics and information technology, a system can be represented through a model using various methods, including mathematical equations, symbols, data, or a combination (Mansouri, Sivaram, Savoie, & Gani, 2025). In a general perspective, considering the human brain as a biological dynamic system, a functional causal network *in silico* models can effectively illustrate. Recognizing the therapeutic implications of brain injuries and analyzing activation data from functional neuroimaging necessitates precise modeling of the connections between cognitive functions and network areas (Labatut, Pastor, & Ruff, 2002). From another perspective, Artificial Intelligence (AI) improves diagnosis, treatment planning, and outcome prediction, transforming brain tumor management (Abdel Razek et al., 2021).

AI-based models can enhance tumor segmentation, grading, and molecular characterization of diseases (Luo, Pan, Mo, Mao, & Zou, 2023), enabling personalized therapy (Cè et al., 2023). The combination of machine learning with three-dimensional (3D) bioprinting, a technical process that utilizes layers of bio-inks to fabricate tissue-like structures, has shown success in mimicking tumor characteristics and predicting therapy outcomes. This combination may lead to the incorporation of immunotherapy and anti-angiogenic agents as feasible treatment alternatives (Min Tang et al., 2024). AI-based surgical planning techniques are also beneficial in identifying the best resection borders while preserving the patient's quality of life.

In adults, the most prevalent and deadly malignant brain tumor in the central nervous system (CNS) is glioblastoma (GB). Gliomas are among the most prevalent types of primary brain tumors (Wesseling & Capper, 2018). Gliomas are among the most common categories of principal cerebral neoplasms (UCSF, 2024). GB is a subtype of glioma (University of Texas). GB’s aggressiveness mainly stems from its extensive migration into the surrounding brain tissue, a process affected by astrocytes, the glial cells of the CNS with a supportive function. Worldwide, the yearly incidence is

approximately six cases per 100,000 people (Bray et al., 2024). The risk factors encompass age, familial history, and radiation exposure (Ostrom et al., 2023). The main clinical signs of GBs are convulsions, high intracranial pressure, and neurological and cognitive impairment. The most common imaging techniques used for identifying GBs are magnetic resonance imaging (MRI) (Shukla et al., 2017) as well as computed tomography (CT) (Alnawafleh, Radzi, Alshipli, Oglat, & Alflahat, 2024); these are augmented by needle or stereotactic biopsies to confirm the diagnosis (Luo et al., 2023). GB is frequently managed with optimum safe resection, radiotherapy, and chemotherapy utilizing temozolomide (TMZ). On the other hand, chemoradiation and/or surgery may be used to treat lower-grade GBs. Numerous clinical trials are being conducted to assess the effectiveness of immunotherapy (Cè et al., 2023). Microscopy techniques are used to visualize GB cell cultures (Krieger et al., 2020) or GB patient-derived cells (PDCs) (Y. Fan et al., 2018). Interacting GB cells exhibit behaviors such as migration, invasion (P. Sharma, Aaroe, Liang, & Puduvalli, 2023), and proliferation (H. L. Li et al., 2020). Additionally, the morphology of GB cells can vary among particular GB cell lines (Diao et al., 2019).

The application of AI in spatial technologies is enhancing our understanding of the tumor microenvironment (TME), which is beneficial for the development of personalized treatments (Walsh & Quail, 2023). Together, the numerous cellular and acellular elements that make up the TME promote tumor growth, metastasis, invasion, and responsiveness to treatment (Elhanani, Ben-Uri, & Keren, 2023). In other words, observing the TME's cellular heterogeneity and spatial structure is becoming increasingly essential to elucidate the mechanisms that drive disease progression and treatment response. More research is required to realize AI's potential in GB management fully. Despite these gains, challenges persist in standardizing AI models to enhance generalizability and interpretability (Luo et al., 2023). The use of AI in GB research in 3D microenvironments has grown recently, yielding encouraging results despite the technology's limitations and challenges (Walsh & Quail, 2023).

Time series data refers to data acquired at regular intervals across time. Time series data often exist in all fields, and they are accessible in finance, meteorology, signal processing, healthcare, and economics. Weather data, market values, and sales figures are some examples of time series data. Time series data analysis enables the forecasting

of conditions and the prediction of future situations by utilizing patterns and values derived from historical data through a prediction methodology. Certain researchers compare LSTM with Transformer models for predicting the future tendency of data (Sonata & Heryadi, 2024). In our study, we captured image cell data on 3D multiplex cell culture TME platforms in a time-series format and predicted the future behavior of cells.

Hybrid predictive models are widely used in engineering literature for the remaining usable life (RUL) of lithium-ion batteries, utilizing data decomposition, transformers, and DNNs to enhance accuracy in particular scenarios with limited historical data. Scientists utilized complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) to analyze sequential data and predict the performance of capacity regeneration components with higher accuracy. They use DNNs to determine how the world's conditions evolve over time in the context of environment and engineering. The model operates accurately with only 25% to 30% of the data, yielding errors of 0.0208 and 0.0337, respectively. Therefore, data decomposition, transformers, and DNN approaches are combined to develop a hybrid model, and scientists have obtained accurate results (Cai et al., 2023).

Scientists have evaluated the effectiveness of transformers in comparison to other neural networks for processing audio data. In a pertinent study, scientists utilize ten distinct music genres and an initial modification to facilitate data comprehension. The neural network structures are examined, and the results are contrasted across various scenarios. The study also examines how configurations and optimization affect the performance of transformers (Ali, 2024).

In the literature, some studies utilize transformers for data decomposition, while deep neural networks (DNNs) may also be combined to develop a machine learning (ML) model (Cai et al., 2023). Hybrid models combine particular machine learning approaches, including the integration of DL methods with convolutional networks. The objective is to provide more accurate predictions that can be useful for clinicians, patients, and researchers in the context of 3D cell culture multiplex platforms. The following image (Figure 1.1), concerning 3D confocal microscopy images of GFP-tagged I-NHA human brain astrocytes and RFP-tagged human brain U87 GB cells, was

obtained using these ECM multiplex platforms (Ismayilzada, 2024). The advantage of such platforms is that they enable the simultaneous capture of biological replicates under the same conditions, providing more input data for ML algorithms and increased accuracy in predictions.

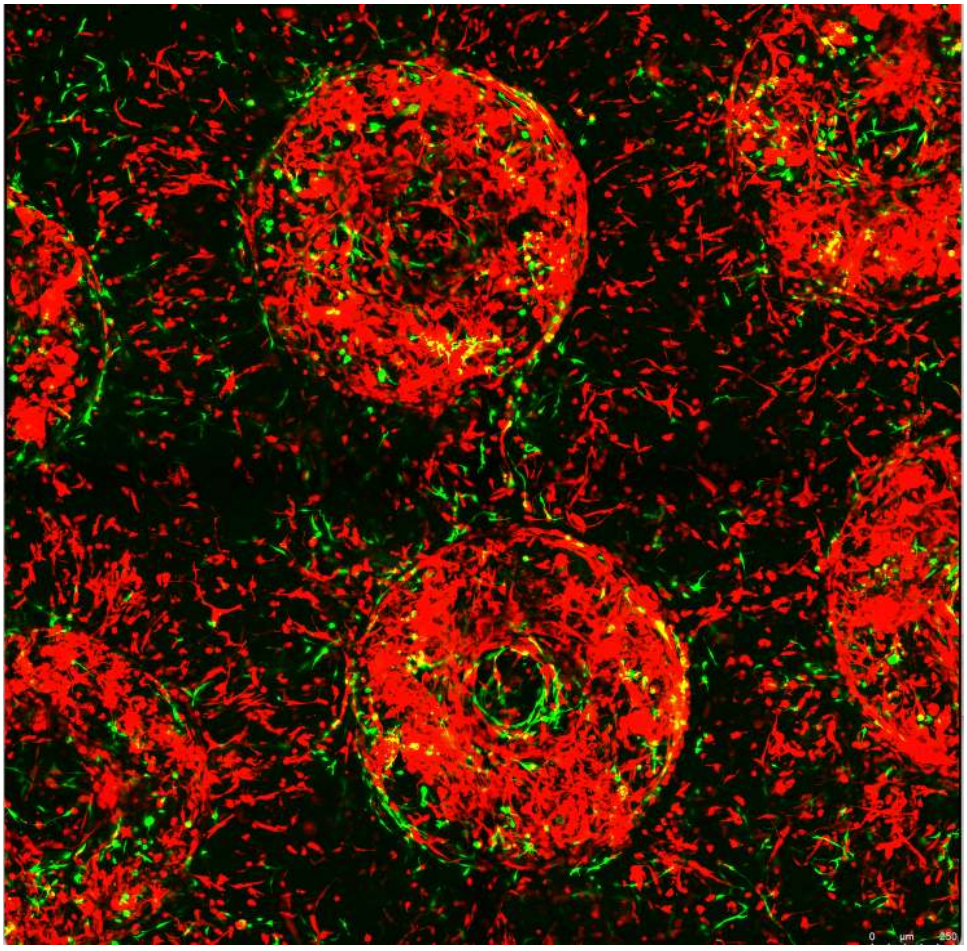

Figure 1.1: 3D confocal microscopy images of RFP-tagged human brain U87 GB cells and GFP-tagged I-NHA human brain astrocytes in direct co-culture configuration obtained with multiplex ECM

In the literature, hybrid models are also used in the field of neuroscience. A hybrid DL transformer model for predicting epileptic seizures is a promising tool for clinicians to use to evaluate patients. The proposed hybrid framework for automated epilepsy screening utilizes EEG data to evaluate seizures, combining a feature extraction module with a deep transformer model to identify hidden features that enable accurate predictions. The framework was tested using the CHB-MIT database and achieved an

average sensitivity of 95.2% and a false positive rate of 0.02, indicating that the results are accurate enough for predictions (Maddineni, Janapati, Kosana, & Teeparthi, 2022). Another study in the field of neuroscience reveals a hybrid DL model for diagnosing obstructive sleep apnea by integrating CNN and Transformer methodologies for radar-based detection of apnea-hypopnea events (J. W. Choi et al., 2024).

Transformers are utilized for treating neurodegenerative disorders, including Alzheimer's disease (AD), while using 3D MRI scans (Hu, Wang, Jin, & Hou, 2023). Scientists have introduced a novel Vision Transformer-integrated Convolutional Neural Network (VECNN) that applies 3D MRI to improve diagnostic accuracy. The model utilized the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, comprising 2.248 MRI images. It could distinguish between AD and healthy controls, as well as moderate cognitive impairment, with an accuracy of 92.14%, precision of 86.84%, sensitivity of 93.27%, and specificity of 89.95%. VECNN is a promising, non-invasive, cost-effective, and objective diagnostic tool that could lead to earlier disease detection and individualized treatment for AD (Zhao et al., 2024).

In the field of BT, transformers are already being used as an ML model in the following ways: prediction route enhancement of pharmacological combinations, predicting voxel-based volume calculations for dosage and radiation outcomes, classifying brain metastases into primary organ locations, predicting molecular expressions, categorizing and grading brain cancer, and segmenting brain tumors (Lan, Zou, Qin, & Zhu, 2023). The following figure (Figure 1.2) summarizes the utility of transformers in brain tumor research.

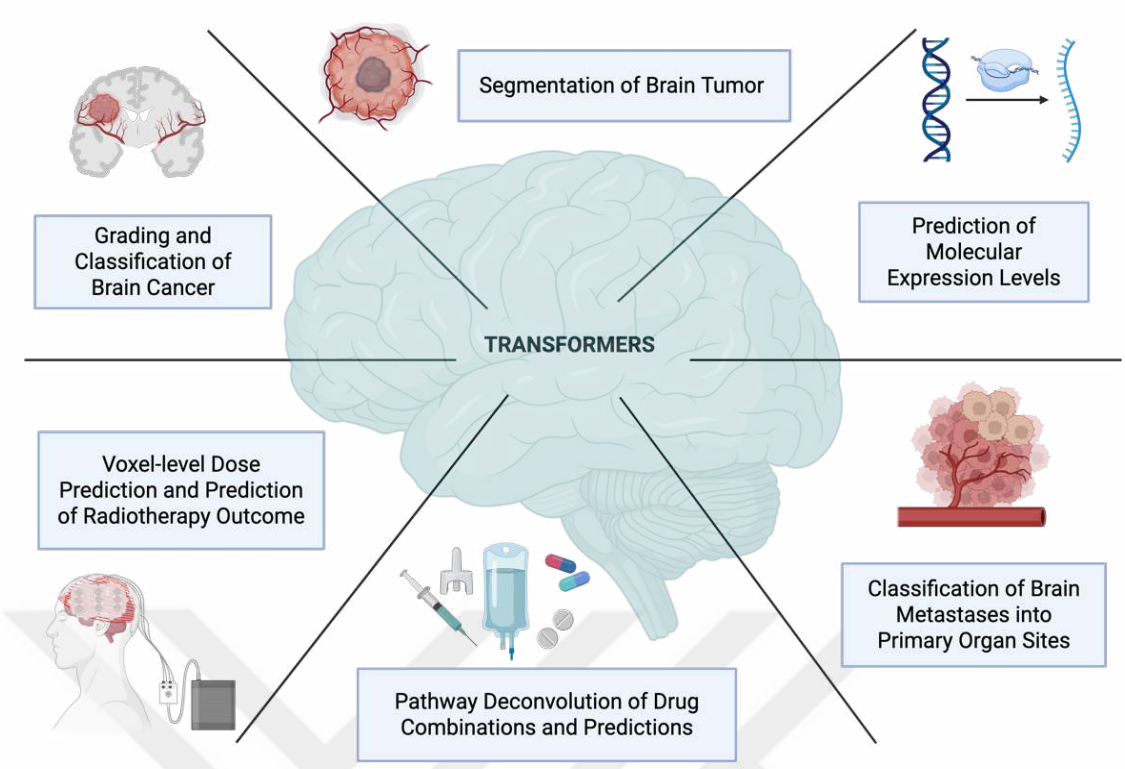

Figure 1.2: Usage of transformers as an ML model in BT research (Lan et al., 2023)

All AI-based studies related to GB are valuable, promising, and helpful for patients and clinicians in facilitating their decision-making. Although AI applications range from disease diagnosis and severity prediction to prognosis and treatment, as well as post-operative care, this article systematically considers AI methods in the use of 3D *in vitro* technologies (Luo et al., 2023).

This chapter summarizes how AI is currently used to explore GB with 3D imaging technologies. It emphasizes the distinct advantages of this approach in modeling GB cell interactions, predicting therapeutic responses, and personalizing treatments. Although this field shows potential, it is confronted with various obstacles. These include the intricate nature of data integration, the limited accuracy of *in vitro* models in representing biological processes, and the necessity for more uniform 3D platforms. It is essential to consider these limitations to apply AI-based findings in clinical practice effectively. This section outlines a roadmap for future research directions that could transform how GB is studied and treated at the nexus of AI and 3D technologies.

1.1 AI and its general concepts

The potential of AI to analyze extensive datasets and identify patterns that exceed human comprehension has made it a powerful tool in neuroscience. Within the realm of GB, AI has shown promise in enhancing diagnostic precision, predicting patient outcomes, and identifying new avenues for therapeutic intervention. AI has become a valuable tool in biomedical research, enabling the analysis of large datasets, pattern identification, and outcome prediction (Al-Rahbi, Al-Mahrouqi, & Al-Saadi, 2024).

The connection between AI and neuroscience has evolved through a significant history of collaboration. Recent advancements in neuroscience, along with substantial enhancements in computational strength over the last several decades, have led to the emergence of a new generation of *in silico* neural networks, computational models designed to emulate the function and structure of the human brain. These biology-inspired architectures emulate biological neural networks by utilizing layers of interconnected artificial neurons that compute and transmit information through synaptic communication, similar to the brain.

Contemporary AI systems exhibit capabilities that mirror the sophisticated perceptual and cognitive functions of biological systems, such as object recognition and decision-making processes. Furthermore, the application of AI in neuroscience research is on the rise, significantly enhancing our comprehension of brain functions (Macpherson et al., 2021).

Some research examines brain-inspired learning representations in artificial neural networks, incorporating biological principles such as synaptic plasticity. This analysis examines the benefits and obstacles of this methodology, emphasizing prospective research directions in understanding the nature of intelligence while underscoring potential advantages and challenges in this rapidly evolving domain (Schmidgall et al., 2024).

Neuroscience and AI have a lengthy, interconnected history. It has been contended that advancements in neuroscience have been, and continue to be, pivotal in driving the evolution of new approaches in AI technology. Analyzing these previous claims uncovers a more complex narrative, in which AI scientists have been subtly inspired by

the brain; however, concepts predominantly originated from the latter (Gershman, 2024).

Human consciousness is a state of awareness generated by brain structures and neurotransmitter systems. The discussion on artificial consciousness continues, with proponents of strong AI comparing it to cognitive states. At the same time, advocates of weak AI argue that machines merely replicate thought without achieving genuine awareness. Some studies rigorously analyze neuroscience-based frameworks for artificial consciousness, highlighting the challenges and prospects in developing computers that can replicate conscious states. They also tackle ethical and philosophical issues (Sadegh-Zadeh & Bahrami, 2025).

Neuromorphic computers, which replicate brain architecture and employ event-driven processing, are gaining attention due to the increasing power requirements in contemporary computing. These systems, comprising brain-inspired neural networks, artificial neural networks, and hybrid circuits, utilize neuromorphic sensors as input mechanisms. Three distinct neuromorphic computers, such as SpiNNaker, TrueNorth, and Loihi, alongside a dynamic vision sensor-based electronic retina, exemplify the potential for low-power, cost-effective, and adaptive solutions to social and scientific challenges in the future (Enuganti, Sen Bhattacharya, Serrano Gotarredona, & Rhodes, 2025).

Implementing deep networks and associated techniques in AI systems has led to significant transformations in mathematics, computer science, and various domains, including neuroscience. Researchers have demonstrated a higher level of effectiveness compared to earlier established methods in key domains, thanks to advancements in AI research, image processing, robotics, natural language processing models, and manufacturing, particularly in the realm of complex games. Practical applications are extensively utilized in areas such as computer vision and speech and text recognition, with significant initiatives undertaken across multiple domains (Ullman, 2019). The following table (Table 1.1) illustrates the primary AI algorithms, including computer vision algorithms commonly employed for image analysis and prediction (Badai, Bu, & Zhang, 2020).

Table 1.1: AI Applications: AI algorithms, primarily computer vision algorithms, which are frequently used for image analysis and prediction.

	Artificial Intelligence Models
Vision Related Algorithms	Computer Vision Models
	Natural Language Processing Algorithms (NLP)
Machine Learning and Deep Learning Related Algorithms	Reinforcement Learning Models
	Generative Models
	Graph Neural Networks
	Time Series Forecasting Models
	Anomaly Detection Models
	Multi-modal Models
	Diffusion Models
Robotics Related Algorithms	Transformers for Various Applications
	Intelligent Robots and Vehicles Applications

This section provides a roadmap for leveraging AI to develop more effective GB treatments by highlighting recent advancements and identifying future challenges, with a focus on AI applications in 3D imaging technologies (Bai et al., 2024). We observe prospective roles of AI in the management of GB, highlighting its utility in diagnosis, treatment formulation, prognosis assessment, and pharmaceutical development. AI algorithms have shown considerable potential in providing therapeutics through imaging analysis, radiomics, and tumor segmentation. These technologies may enable non-invasive molecular profiling and the early identification of GB (A. Rončević, N. Koruga, A. Soldo Koruga, & R. Rončević, 2025).

Combining AI and 3D imaging technologies can enhance our comprehension of the GB TME, optimize treatment approaches, and expedite drug discovery (Issa, Stathias, Schürer, & Dakshanamurthy, 2021). Neuroscientists must bridge the gap between emerging technology and experimental research through multidisciplinary studies to elucidate the mechanisms underlying neurological disorders and offer new insights and treatments. Although advancements in AI and its applications have exceeded expectations, a significant divide persists between AI and human intelligence (J. Fan,

Fang, Wu, Guo, & Dai, 2020). Progress in 3D culture methodologies has resulted in the creation of more authentic tumor models, including bioprinting, which can promote the advancement of more effective pharmacological treatments for GB, a common and fatal brain cancer. We observe studies delineating essential elements for 3D bio-printed GB models, encompassing the examination of biomaterials and crosslinking techniques, and assessing their incorporation into GB-on-a-chip systems. The proposal outlines a Quality-by-Design framework for design and examines future prospects, including 4D bioprinting and machine learning methodologies, as well as existing gaps that must be addressed to progress further in the area (Branco, Cunha, Mendes, Sousa, & Vitorino, 2025).

AI is a broad term encompassing any computer-implemented work that requires human intelligence, including systems based on explicit rules and computational methods lacking rigidly defined rules. Two core ideas are reviewed to enhance comprehension of today's capabilities and potential AI applications: Machine Learning (ML) and Deep Learning (DL) (Parekh & Jacobs, 2019).

Several studies have analyzed the application of DL in glioma segmentation using MRI data, emphasizing the integration of deep learning with the effectiveness of convolutional neural networks in tumor segmentation. They analyze clinical applications and the advancement of deep learning-based segmentation, focusing on data quality and prospective research avenues. This venture highlights the importance of addressing tumor heterogeneity and the judicious application of DL-driven healthcare technology (Ghadimi et al., 2025). Another study uses a 2D U-Net model to precisely separate gliomas in MRI scans, showcasing its applicability in practical neuroimaging contexts. The model attained an 81% Intersection over Union and a 65.5% Dice score (Amri, Ben Slama, Mbarki, Selmi, & Trabelsi, 2025).

Some imaging techniques are used during the treatment process of GB. The Stupp procedure, a standard treatment for GB, aims to improve patient outcomes by evaluating various medications in optimal combinations. Medical gadgets augment the protocol's effectiveness, with tumor-treating fields demonstrating superior patient survival rates. Future developments encompass advanced imaging techniques, individualized medicines, and nanotechnology-enabled devices for targeted therapeutic delivery to

tumor locations. These discoveries generate optimism for GB sufferers and healthcare professionals (Mnich et al., 2025). Another study examines the use of the 3D-Slicer multimodal image fusion technique in the invasive treatment of GB, aiming to improve preoperative planning, mitigate surgical issues, and reduce postoperative challenges (Liu & Wang, 2024).

Technological advancements in diagnostics, treatment, and translational research facilitate precision radiotherapy for glioblastoma. Biologically informed multimodal imaging can tackle treatment resistance, offering safety and possible advantages. Initial clinical trials have identified imaging biomarkers for personalized radiation therapy. Future prospects encompass response-adaptive radiation utilizing biomarkers and innovative systemic medicines, leading to enhanced glioblastoma outcomes and individualized patient care (Breen, Aryal, Cao, & Kim, 2024).

A study on data presents a brain tumor database of 23,049 samples, each containing four types of MRI images. Ground truth is provided for each sample, aiding interpretation, and improving network training. The dataset was derived from 438 samples extracted from the original dataset provided by the BraTS 2022 Challenge. The diverse image quality reflects the inequities in clinical practices between hospitals. This dataset is suitable for advancing automated segmentation techniques for patients with newly diagnosed brain tumors (Solak et al., 2025).

Malignant glioma is characterized by the unrestrained proliferation of cells in the spinal cord and brain, leading to neurological deficits, convulsions, hormonal imbalances, and venous thromboembolism. Medical examinations, including MRI scans, CT scans, biopsies, and electroencephalograms, facilitate early detection of diseases. Deep learning models are optimal for disease prediction but necessitate substantial memory and storage capacity. A methodology based on distributed and privacy-preserving horizontal federated learning has been developed to address these challenges. The model attained the highest accuracy with five clients' architectures, achieving 99.76% for IID and 99.71% for non-IID distributions, respectively (P. Sharma et al., 2023).

A summary of the key ideas in AI is provided in the following graph (Figure *1.3*). There are definitions for ML and DL, two essential terms of AI (R. Y.

Choi, Coyner, Kalpathy-Cramer, Chiang, & Campbell, 2020). CNN refers to convolutional neural networks (Nielsen, 2015), RNN (J. Wang, Zhang, Yang, Yeh, & Wang, 2022) for recurrent neural networks, DNN for deep neural networks (Luo et al., 2023), showing that developing a DNN for cell counting and classification is feasible (Jiang et al., 2021).

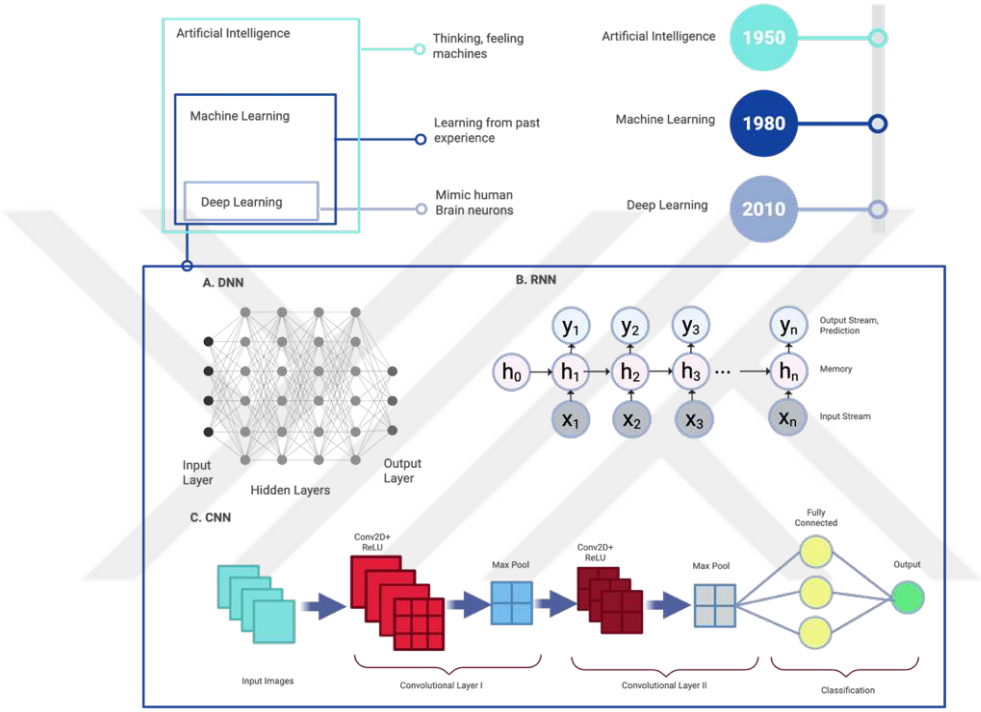

Figure 1.3: Brief outline of the fundamental concepts in artificial intelligence. Definitions arise for two fundamental terms: machine learning and deep learning. CNN denotes convolutional neural networks, RNN signifies recurrent neural networks, DNN represents deep neural networks (Luo et al., 2023)

Certain works investigate the utilization of AI in GB research, highlighting the discovery of new biomarkers and therapeutic targets. These results indicate an increasing interest in AI-based approaches for biomarker and drug research in GB. The findings suggest encouraging preliminary outcomes; nevertheless, additional extensive and methodologically sound research is required to confirm practical applications and establish standardized protocols for enhanced reproducibility and clinical implementation (Conte et al., 2025).

1.2 Machine learning

A specific category of AI, known as machine learning, enables a computer to acquire knowledge from existing training data, rather than relying exclusively on explicit programming, to predict new data. It emphasizes hypothesis generation over hypothesis testing (Vicini et al., 2022). A predictive model based on ML is constructed in various stages: determining the prediction, creating, or maintaining the pertinent datasets, building the model, and evaluating the forecast's effectiveness. Unsupervised and supervised learning are the two primary categories into which machine learning techniques are classified. Unsupervised learning utilizes features to identify categories, or natural groupings, in data without labels (Cellina et al., 2022).

On the other hand, datasets with labeled data (classes) are used in supervised learning to train or construct models that link characteristics to categories. Artificial neural networks (ANN), a subset of machine learning approaches, emulate the human brain's neural network to process information, hence generating simplified models and forming separate networks based on various connection methodologies (Zhongheng Zhang, 2016). ANN functions via numerous interconnected nodes (neurons), each embodying a specific output function (termed an activation function). The relation between two nodes signifies the weighted value of the signal traversing that link (referred to as weight), analogous to the ANN's memory (Agatonovic-Kustrin & Beresford, 2000). The excitation function, weights, and connecting method affect the network's output. In basic terms, the network typically serves as an approximation of a function, algorithm, or logical methodology (Luo et al., 2023).

The following table (Table 1.2) demonstrates the most common ML algorithms based on three subtypes: Supervised, unsupervised, and reinforcement learning.

Table 1.2: Machine learning algorithms, illustration of the most prevalent algorithms classified into three subcategories: supervised, unsupervised, and reinforcement learning

Machine Learning Algorithms		
Supervised Learning	**Unsupervised Learning**	**Reinforcement Learning (RL)**
Support-Vector Machines	K-means Clustering	Q-Learning
Linear Regression	Hierarchal Clustering	SARSA
Logistic Regression	Anomaly Detection	Deep Q Network (DQN)
Naive Bayes	Neural Networks	Deep Deterministic Policy Gradient (DDPG)
Linear Discriminant Analysis	Principle Component Analysis (PCA)	TRPO (Trust Region Policy Opt.)
Decision Trees	Independent Component Analysis	PPO (Proximal Policy Opt.)
K-nearest Neighbor Algorithm	Apriori Algorithm	Genetic Algorithm RL
Neural Networks (Multilayer Perceptron)		

1.3 Evolution of Artificial Intelligence: from basic machine learning to neural networks

A subset of neural network-based ML known as DL can be identified by its many layers, especially pertinent to images and videos. The ability to convert the provided data into an abstract and combined representation may be learned at each level. The procedure establishes where features at various levels should be placed. Image analysis has been substantially improved by DL, which aims to mimic neuronal function and interaction within a multilayer neural network. This approach, which involves training on extensive datasets of digital images, enhances object recognition in images beyond the capabilities of human cognition. Multiple variables and layers define a DNN, an ANN. DL encompasses the domain of machine learning that utilizes DNNs, with a particular focus on learning algorithms and substantial quantities of training data (Luo et al., 2023).

DL primarily consists of Long Short-Term Memory (LSTM) networks, CNN, and recurrent neural networks (RNN). CNN represents the most recent advancement in imaging analysis technology, utilizing a specialized ANN architecture that employs mathematical convolution operations to transmit information across several layers, illustrating the interconnected neuronal patterns and states in the human visual cortex (Luo et al., 2023).

In the field of medicine, solutions for numerous diseases have been discovered through the systematic gathering of data from patients and the analysis of their similarities and

variations. Given the significant increase in data volume and the variety of application domains, we have reached a stage where automating the transition from data to knowledge is essential (Alpaydin, 2021). The large datasets generated from the 3D GB models were analyzed using supervised and unsupervised machine learning algorithms (Luo et al., 2023).

1.3.1 Unsupervised learning

Unsupervised learning (UL) approaches rely solely on the intrinsic structure of unlabeled data, exhibiting self-organization to uncover hidden patterns within the data (Jung, 2022). In this category of machine learning, the ML model is supplied with a basic dataset and instructed to extract insights from it (Müller & Guido, 2016). They are employed in tasks such as algorithms that involve clustering, which aims to partition information into groups based on specific feature attributes, and association rules focused on uncovering connections between data points (Cè et al., 2023).

UL methods, including clustering algorithms and principal component analysis (PCA), were utilized to investigate concealed patterns in tumor cell dynamics and their interactions with the (Akbari et al., 2021).

1.3.2 Supervised learning

The core process of ML that requires human supervision for training is called supervised learning (SL). These algorithms are appropriate for general categorization tasks because they mimic the human cognitive mechanism of learning by exemplification (Shalev-Shwartz & Ben-David, 2014). Experts in software development must annotate input data points with appropriate label values in a training dataset consisting of input-output pairs (Sidey-Gibbons & Sidey-Gibbons, 2019). By simulating human annotators who gain knowledge via experience, the program maps potential links among data to test theoretical hypotheses (Cè et al., 2023). Random forests, support vector machines (SVMs), and DL models exemplify SL methodologies (Müller & Guido, 2016). These algorithms predict treatment responses, ascertain biomarkers, and classify cell types within the TME.

Support vector machines (SVMs) can be used to predict time series. Scientists frequently employ time series prediction methods in fields such as weather prediction, load prediction on an electric grid, and reliability prediction. These applications often involve complex and intricate foundational system models. Conventional linear estimating techniques could be inadequate, requiring sophisticated algorithms. Some researchers examine the application of SVMs, emphasizing their efficacy in forecasting under conditions of nonlinearity and non-stationarity in system dynamics. In specific situations, SVMs may be more accurate than other non-linear methods, including neural networks (Sapankevych & Sankar, 2009).

1.3.3 Semi-supervised learning

Between unsupervised and supervised learning, semi-supervised learning (SSL) is an ML paradigm that primarily utilizes unlabeled data, along with a limited quantity of labeled data (Cè et al., 2023). SSL denotes a subset of supervised learning tasks and techniques that utilize unlabeled data for training, typically involving a small amount of labeled data in conjunction with a large volume of unlabeled data. SSL occupies a position between UL, which operates without any labeled training data, and SL, which relies entirely on labeled training data. A significant number of machine learning experts have discovered that using unlabeled data alongside a limited set of labeled data can lead to substantial enhancements in learning accuracy (Learning, 2006).

1.3.4 Self-supervised learning

Self-supervised learning refers to an approach in machine learning where the fundamental principle is the autonomous generation of a supervisory signal to address a specific problem. Self-supervised learning resembles unsupervised learning due to the absence of labels; nonetheless, it addresses challenges often associated with supervised learning. This type of ML addresses the problem of limited data availability by utilizing the enormous amount of accessible but unlabeled data to train accurate classifiers, due to the challenges of maintaining comprehensive control of all data in a large-scale venture. Occasionally, specific neural networks, such as autoencoders, are defined as self-supervised learning mechanisms (Cè et al., 2023).

1.4 Preprocessing and data normalization

In numerous contemporary applications, such as bioinformatics and computer vision, samples exhibit various feature representations derived from diverse data sources (Gönen & Margolin, 2014). Before conducting AI analysis, the datasets obtained from the 3D models underwent preprocessing. The imaging and molecular data underwent a thorough cleaning, normalization, and annotation process to ensure high quality and consistency across various experiments (Way et al., 2017). The concept of data augmentation involves techniques for creating iterative optimization or sampling algorithms by integrating unseen data or latent variables (Dulam, Gade, & Gosukonda, 2023). Enhancing model performance through data augmentation enables the expansion of limited datasets, thereby leveraging the potential of big data (Shorten & Khoshgoftaar, 2019). Methods for data augmentation were employed to increase the size of the datasets used to train AI models (Van Dyk & Meng, 2001).

1.5 Features

ML generally assesses qualitative characteristics, such as variations in signal intensity, which are measurable and quantifiable. Visual qualities or metrics that can be readily computed or quantified by a computer, in contrast to subjective reference standards. A comparison is typically made between features automatically extracted via deep learning techniques and features manually generated by an expert (Müller & Guido, 2016). First-level features include attributes easily deduced from a qualitative standpoint, such as shape, volume, and intensity signal measurements. On the contrary, second-order traits may seem clinically uninformative to the interpreter because they are either automatically extracted by DL networks or are created from the combination of first-level data. Therefore, in the radiomic workflow, choosing robust and reproducible features is essential (Sansone et al., 2023). An ML paradigm known as self-supervised learning involves the basic concept of autonomously generating a supervisory signal to address a specific issue. Self-supervised learning closely resembles UL due to the absence of labels; nonetheless, it seeks to address challenges often associated with supervised learning. Due to the inability to ensure sufficient oversight for all data required by a major project, this ML tackles the issue of limited data availability by

utilizing the enormous quantity of readily available yet unannotated data for training accurate classifiers. Certain neural networks, including autoencoders, are occasionally characterized as self-supervised learning systems (Cè et al., 2023).

1.6 AI-based predictions and simulations

AI models were trained to predict drug efficacy, tumor progression, and immune responses within the 3D models. The models were validated using a subset of experimental data to ensure the projections aligned with biological outcomes. In addition, AI-based simulations were conducted to model potential therapeutic strategies, specifically combination therapies targeting multiple GB TME components (Wu et al., 2022).

1.7 Validation and model refinement

The verification of the algorithms and software is established through the iterative metamorphic testing methodology, targeting the non-testable attributes intrinsic to these systems. A methodical machine learning methodology is introduced to create test oracles, to guarantee sufficient test coverage criteria (Ding, Hu, & Gudivada, 2021).

1.7.1 Verification of AI predictions in GB cultures

The accuracy of the AI-generated predictions should be confirmed by comparing them to the experimental outcomes obtained from the biomimetic 3D models. The analysis compared anticipated drug reactions, cell proliferation rates, and invasion patterns with real-world observations from treated GB cultures (Luo et al., 2023). In AI-related studies, multiple experimental replicates are employed to guarantee the robustness and reproducibility of AI predictions. The efficacy of the machine learning algorithm is assessed using a stratified N-fold cross-validation technique in conjunction with a confusion matrix (Ding et al., 2021).

1.7.2 Validation in living organisms

When possible, the accuracy of AI models based on 3D cultures was further confirmed by conducting experiments on animal models of GB. This step was crucial in ensuring that the AI-based insights could be applied to biological systems beyond *in vitro* models. The enhanced and preprocessed data are subsequently fed into machine learning models to optimize the model parameters, to minimize the cost function. Simultaneously, implicit data characteristics will be extracted using statistical techniques. The machine learning models are validated during the training process to reduce overfitting. In this case, the trained artificial neural network model excels on the training dataset but does not generalize effectively to novel examples (Zhu et al., 2022).

1.7.3 Model refinement

Model refinement is a crucial process in machine learning, focusing on enhancing the performance and generalization of predictive models. After analyzing the validation outcomes, adjustments were made to the AI algorithms to enhance their accuracy. The process included training models with larger datasets, integrating new experimental data, and fine-tuning model parameters. Hyperparameter optimization methods are crucial for refining models. Hyperparameters, including learning rates and regularization abilities, influence model performance. Techniques such as random search, grid search, evolutionary algorithms, and Bayesian optimization efficiently explore the hyperparameter space to identify optimal configurations. By utilizing regularization and hyperparameter optimization, machine learning practitioners can refine models, strike a balance between complexity and generalization, and ultimately enhance prediction accuracy (Krishnan, Kodamana, & Bhattoo, 2024).

1.8 Analysis of data using statistical methods

Besides ML methods, datasets can be used in data analyses conducted using appropriate statistical methods, including significance testing such as t-tests and ANOVA, as well as the Kruskal-Wallis test, to evaluate the results. It is possible to integrate RNA sequencing data and MRI images and conduct both ML and statistical methods on the datasets (Fathi Kazerooni et al., 2025). Calculations can be performed to assess the

reliability of AI-based predictions and experimental findings using confidence intervals and p-values (Min Tang et al., 2024).

1.9 Voxel-based volume calculations of 3D brain images

In the literature, various techniques for calculating brain volume are observed. Advanced MR neuroimaging, especially Diffusion Tensor Imaging (DTI), is essential for the diagnosis and comprehension of central nervous system (CNS) disorders. Diffusion Tensor Imaging (DTI), a specialized type of Diffusion-Weighted Imaging (DWI), measures water diffusion in tissues and is particularly useful for assessing traumatic brain injury, whereas DWI serves as the benchmark for identifying acute strokes. DTI offers improved data by measuring the anisotropic diffusion of water along axons. Data from DTI is gathered in 3D voxels, where diffusion values constitute a tensor, facilitating the investigation of water diffusion across several directions. This technique evaluates white matter microstructure, axonal integrity, and tract orientation. Essential measurements in DTI encompass fractional anisotropy (FA), mean diffusivity, radial diffusivity, and axial diffusivity. FA is notably responsive to microstructural alterations and can be assessed using diverse methodologies, including regions of interest, whole-brain voxel-based analysis, or tract-based statistics. These measurements enable doctors to detect anomalies that traditional imaging would otherwise overlook, underscoring the significance of DTI as a vital non-invasive tool in clinical practice and research (Ranzenberger, Das, & Snyder, 2025).

A separate investigation assesses brain volume utilizing the atlas-based approach and a low-strength static magnetic field of 0.4 tesla. An understanding of the repeatability of measured data is crucial for both atlas-based and voxel-based morphometry (VBM) techniques in magnetic resonance (MR) brain volumetry. Nonetheless, numerous recent investigations examining the repeatability of brain volume measurements have utilized static magnetic fields ranging from 1 to 4 tesla, and no research has employed a low-strength static magnetic field (Goto et al., 2016).

Another study reveals that a Brain structural analysis in individuals with post-traumatic anosmia (PTA) utilizes voxel-based morphometry (VBM) and surface-based

morphometry (SBM) to evaluate gray matter (GM) characteristics. This investigation examines neuroanatomical measurements of gray matter and their functional correlates in patients with post-traumatic amnesia, elucidating the structural alterations associated with this condition (Rezaeyan, Asadi, Kamrava, & Zare-Sadeghi, 2023).

Voxel-based methods are present *in silico* research. High-resolution iterative 3D positron emission tomography (PET) image reconstruction necessitates efficient forward-backward projectors to reduce calculation time. These projectors are characterized by a system response matrix (SRM) that includes billions of elements, delineating voxel contributions to lines-of-response (LORs). The efficacy of algorithms such as maximum-likelihood expectation-maximization (MLEM) is impeded by high memory demand and floating-point calculations. The application of symmetries in the SRM facilitates substantial compression, as evidenced by the PET REconstruction Software Toolkit (PRESTO), achieving compression factors above 300 through discrete polar symmetries. Nonetheless, previous implementations encountered inefficiencies attributable to suboptimal memory access patterns. This investigation introduces a symmetry-based assessment to restructure image data and LOR projection data for enhanced vectorization, in conjunction with simultaneous multi-threading (SMT), to achieve speed improvements exceeding 11 times without SMT and beyond 100 times with SMT, all while preserving numerical precision (Scheins, Vahedipour, Pietrzyk, & Shah, 2015).

A separate study reveals findings by examining the intra- and inter-scanner variability of an MRI-based volumetric technique utilizing SPM5 algorithms. Imaging of a healthy volunteer was performed using six distinct MRI scanners to ascertain the coefficients of variation (CV) for multiple brain areas. Intra-scanner CVs ranged from 0.50% to 4.4%, with a median of 0.89%, but inter-scanner CVs varied from 0.66% to 14.7%, with a median of 4.74%. The aggregated variability was marginally elevated (0.87-15.1%, median 4.80%). Substantial volume change thresholds were identified as 1.4% for the identical scanner and 10.5% for disparate scanners. These findings are advantageous for volumetric analyses in neurological disorders and for tracking atrophy patterns or therapy outcomes (Huppertz, Kröll-Seger, Klöppel, Ganz, & Kassubek, 2010).

Mimicking the GB TME in 3D hydrogels and calculating the volume using specific and significant algorithms from 3D images would be valuable for understanding the proliferation and migration behavior of GB cells in the context of this thesis.

1.10 Temozolomide as a drug in GB

Pharmaceutical development is a tedious and progressively expensive endeavor with a seemingly negligible success rate. Identifying the suitable target and developing an effective treatment candidate are deemed the most critical tasks to increase the probability of success in any drug development endeavor. The proliferation of comprehensive genomics datasets and sophisticated AI algorithms is swiftly advancing towards addressing these challenges. The integration of AI and genomics expedites critical elements of the initial stages of drug research (Feizi & Gutierrez, 2022).

TMZ is the primary treatment for glioblastoma patients; however, due to widespread resistance, there is an urgent necessity to produce anticancer medicines for TMZ-resistant GB (Kim et al., 2019). TMZ applications may exist in diverse scenarios, including recurrent GB, GB in geriatric patients, and various glioma classifications such as low- and high-grade gliomas. Recent data demonstrate that TMZ improves survival in older GB patients and those with anaplastic gliomas, exhibiting efficacy with reduced toxicity compared to radiation in high-risk, low-grade gliomas, and yielding results comparable to those of platinum-based chemotherapy in juvenile high-grade gliomas. The successful application transcends glioblastoma, indicating the necessity for discovering biomarkers related to TMZ susceptibility and the development of adjunctive medications to enhance treatment outcomes (Chua, Nafziger, & Leung, 2019). Mimicking the TMZ response in the concept of this thesis on 3D multiplex platforms and observing their results via software calculations, predicting cell behaviors via ML algorithms would be significant and valuable for GB research.

1.11 Ethical considerations related to guidelines

Some studies in the literature integrate 3D technologies and AI to create a robust platform for studying the GB TME. It presents a fresh perspective on comprehending tumor biology and enhancing therapeutic interventions.

AI possesses the capacity to transform healthcare by enhancing diagnostic precision and customizing treatment strategies. Nonetheless, its swift and predominantly unregulated incorporation into medicine raises ethical dilemmas regarding data integrity, patient safety, and effective governance. Transparency is a significant issue, though, as the proprietary nature of numerous large-language models inhibits patients and healthcare providers from comprehending the rationale behind AI-generated results, which may undermine trust. Synthetic datasets produced by AI, while beneficial for enhancing research in data-deficient fields, complicate issues related to data ownership, patient consent, and scientific validity. Moreover, the capacity of generative AI to optimize administrative functions poses a risk of depersonalizing care, increasing the separation between clinicians and patients. Considering these factors, generative AI applications necessitate regulatory and critical examination (Serrano et al., 2024). Fundamental initial measures encompass establishing rigorous standards for data protection and transparency, implementing oversight mechanisms such as institutional review boards to regulate data utilization, and formulating interdisciplinary guidelines involving developers, doctors, and ethicists. By addressing these issues, we can more effectively integrate the deployment of generative AI with the fundamental principles of humanistic healthcare, thereby safeguarding patient safety, autonomy, and trust while leveraging AI's accelerative capabilities (Hasan, Fury, Woo, Kunze, & Ramkumar, 2024).

According to scientists, adherence to ethical guidelines is considered crucial when utilizing AI for analyzing patient-derived biomimetic 3D models (Hofmann, Cohen-Harazi, Maizels, & Koman, 2022). Over the years, the patient data in developing GB models underwent anonymization, and informed consent was acquired for all primary cell lines. AI was utilized for predictive modeling, adhering to ethical standards regarding data privacy and algorithmic transparency.

1.12 Cell culture in 3D microenvironment and disease models

Cell culture has emerged as a crucial tool for elucidating the fundamental biophysical and biomolecular principles underlying cellular assembly into tissues and organs, their functional dynamics, and the disruptions of these functions in pathological conditions. Cell culture is extensively utilized in medical studies, tissue engineering, regeneration medicine, and industrial applications. Despite the predominance of flat, two-dimensional (2D) cell culture, current investigations have evolved into employing 3D structures and more realistic biochemical and biomechanical microenvironments (Duval et al., 2017).

The concept of TME was first introduced in 1979, as researchers have examined how distinctive cells within a tumor interact with one another (Lord, Penney, Sutherland, & Cooper, 1979). Considering GB has a heterogeneous structure, the TME consists not only of tumor cells but also of other cell types, including immune cells, fibroblasts, endothelial cells, and cancer stem cells (CSC) (Gurusamy, Alsayari, Rajasingh, & Rajasingh, 2018). These cells boost tumor growth and spread by interacting with tumor cells and releasing signaling molecules and extracellular matrix (ECM) components. Multiplex profiling assays facilitate the evaluation of cellular behavior in 3D cell culture microenvironments, and they are essential for obtaining novel data in fundamental, translational, and pharmacological research, while enabling the collection of a large amount of data at once on multiplex platforms (Burgstaller, Oehrle, Koch, Lindner, & Eickelberg, 2013).

The following figure (Figure 1.4) illustrates photomasks that enable data capture on multiplex platforms at the micro level (Gurkan et al., 2013). (a) In this example, the tissue structure consists of four distinct types of building blocks, each characterized by distinctive cellular and ECM compositions represented by various color codes. (b) The creation of 3D multilayer tissue prototypes through a layer-by-layer photomasking and alignment technique, utilizing multiple photomasks and various hydrogel solutions that are photo-crosslinked to produce digitally sculpted hydrogels (e.g., incorporating diverse cell types, ECM components, and diffusible factors). The straightforward mask alignment and digital sculpting apparatus consists of: (i) a top cover slip for the fabrication compartment, (ii) a thickness control spacer, (iii) a treated glass cover slip

for immobilizing the sculpted elements, (iv) a fabrication chamber, (v) alignment pins, (vi) a predesigned photomask, and (vii) a mask holder that aligns within the fabrication chamber. Light is transmitted through the aperture of the mask holder. (c-f) Alignment and digital sculpting of individual parts are accomplished using a photomask, which features digitally defined, designed apertures that facilitate the photo-crosslinking of tissue prototypes incrementally. This document demonstrates the fabrication of cubic digitally molded blocks. Radial circular (g) and concentric circular (h) elements can be molded to replicate the cross-section of embedded vasculature within a tissue construct, where a yellow hydrogel encases the three types of elements, microfabricated as concentric circles. (i) The digital sculpting method can also be employed to create intricate structures, such as microscale typescripts.

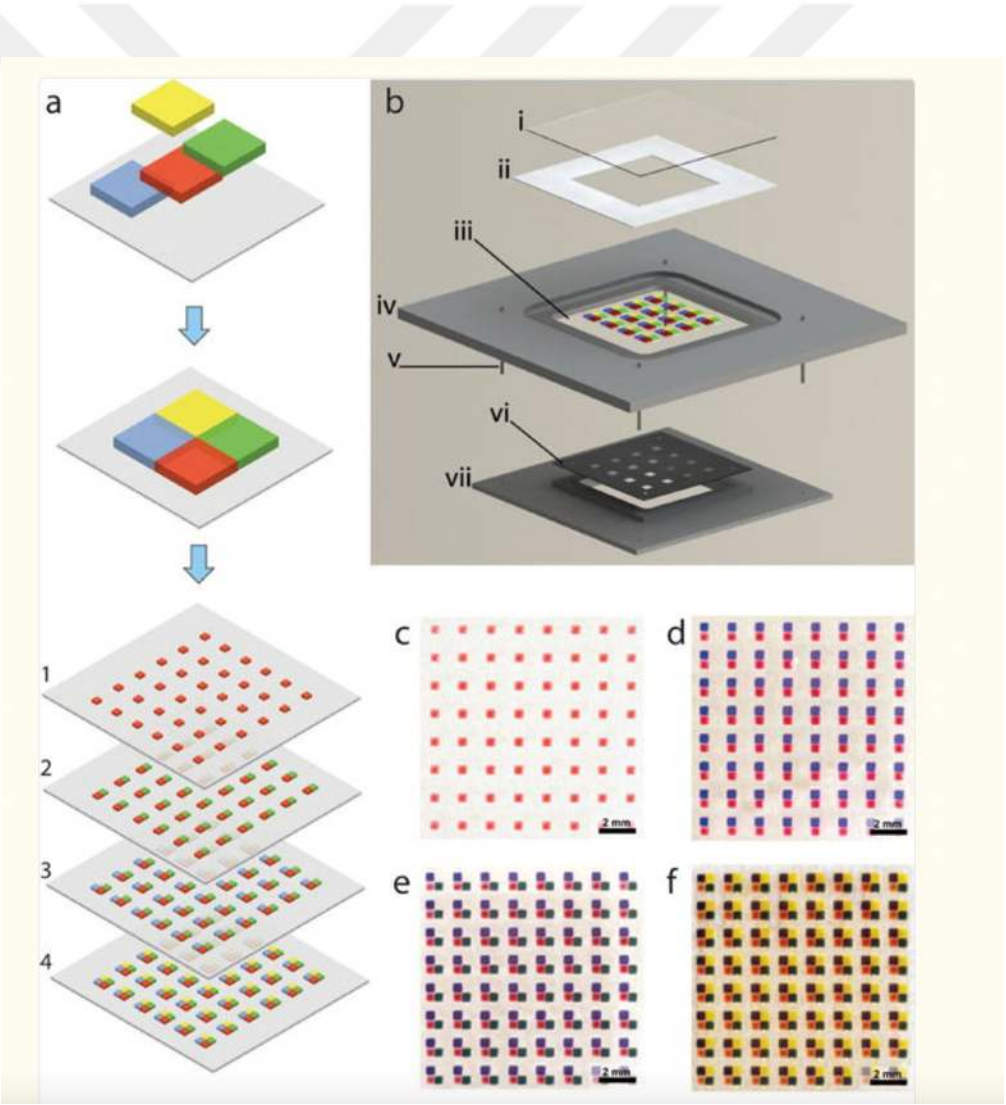

Figure 1.4: Digitally defined spatially heterogeneous microfabrication and tissue manipulation levels (Gurkan et al., 2013)

Gliomas present significant therapeutic challenges due to their complex microenvironments and genetic heterogeneity. The approaches for innovative methodologies that integrate multi-algorithm ML with 3D bioprinting to improve the assessment of glioma treatment response exist in current literature. The bio-printed glioma tissues mimic the molecular characteristics and treatment responses of actual tumors. The GlioML workflow, which utilized an ensemble model and nine algorithms, performed more accurately and effectively than single models in several *in vitro* systems. This approach expanded the range of treatment assessments to include anti-angiogenesis and T cell therapies. The research examined the immunosuppressive and angiogenic tumor microenvironments, identifying potential medications for gliomas and suggesting an improved strategy for treating gliomas, as well as potentially other malignancies (Min Tang et al., 2024).

1.13 Organoids, tumoroids, and spheroids

AI systems integrated with organoids have substantial potential for enhancing our understanding of organogenesis and disease progression, thereby laying the foundation for clinical implementation. AI-integrated organoid systems expedite the swift evaluation of building methodologies, facilitate economical extraction of multiscale picture attributes, and assist in the efficient analysis of multi-omics datasets. Collectively, these features improve the accuracy of preclinical assessments and expand the applicability of organoids in disease modeling, drug development, and personalized medicine. The following (Figure 1.5) summarizes the utility and application areas of the AI-enabled organoids (Bai et al., 2024).

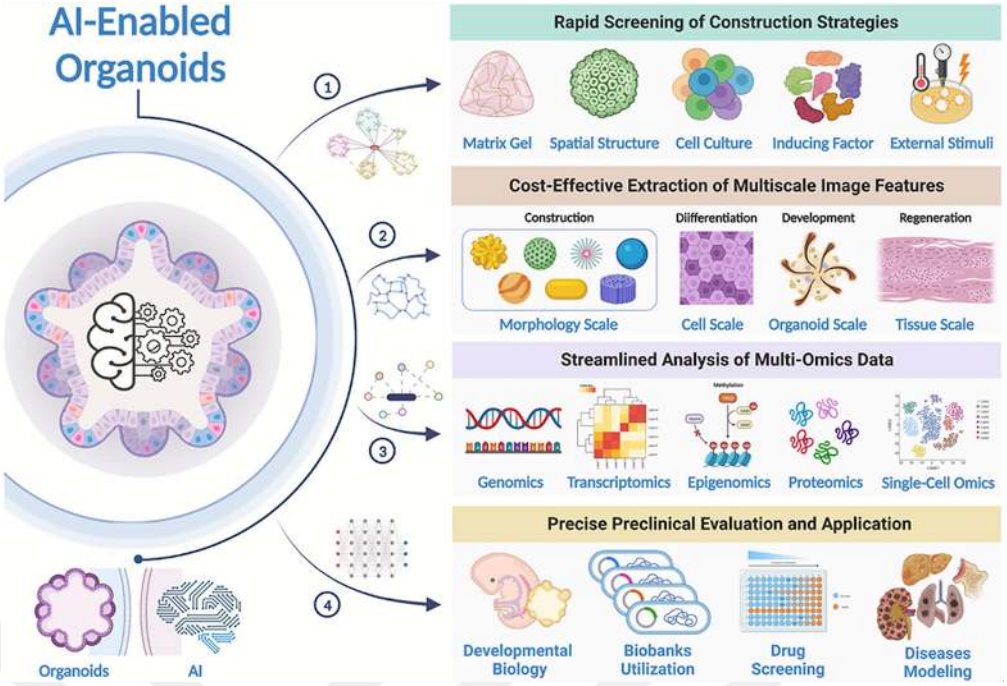

Figure 1.5: Utility and application areas of AI-enabled organoids (Bai et al., 2024)

Organoids have enormous potential for biomedical research, regenerative medicine, and simulating human development and illness. Organoids are 3D cultures made from cells. Organoids have the potential to represent human development and disease since they can be produced from both human stem cells and patient-derived induced pluripotent stem cells. They may also be used in future organ replacement procedures and drug testing. Here, we provide an overview of this rapidly developing topic and discuss how organoid technology may be utilized in future biomedical studies (Lancaster & Knoblich, 2014).

Multicellular tumor spheroids and tumoroids function as superior *in vitro* models that precisely reflect the attributes of the TME. Biomimetic components that replicate the ECM form scaffolds, providing structural support for 3D culture systems, hence promoting the formation of spheroids and tumoroids (C. R. Yang et al., 2022).

1.14 ECM elements

Various elements of the organic ECM can function as substrates in 3D culturing. Natural supports encompass matrigel, hydrogels, rigid substrates, and many synthetic

techniques for 3D matrices, including lyophilization, stereolithography, electrospinning, and microfluidics, each with unique benefits and drawbacks (Habanjar, Diab-Assaf, Caldefie-Chezet, & Delort, 2021).

The morphology of GB cells in 3D ECM may vary depending on the cell type. This distinction was examined in a biomimetic 3D collagen micro-chamber model over a period of three days. Various GB cell types demonstrated unique characteristics: LN229 cells exhibited triangular or diamond shapes, SNB19 cells were elongated, and U251 cells presented irregular shapes with elongated extensions, some of which encroached beyond the micro-chambers by Day 3. Observations were corroborated by high-magnification images and dynamic movies depicting cell morphology and movement (Diao et al., 2019). These behaviors, as demonstrated in cell morphology, are illustrated in the following figure (Figure 1.6).

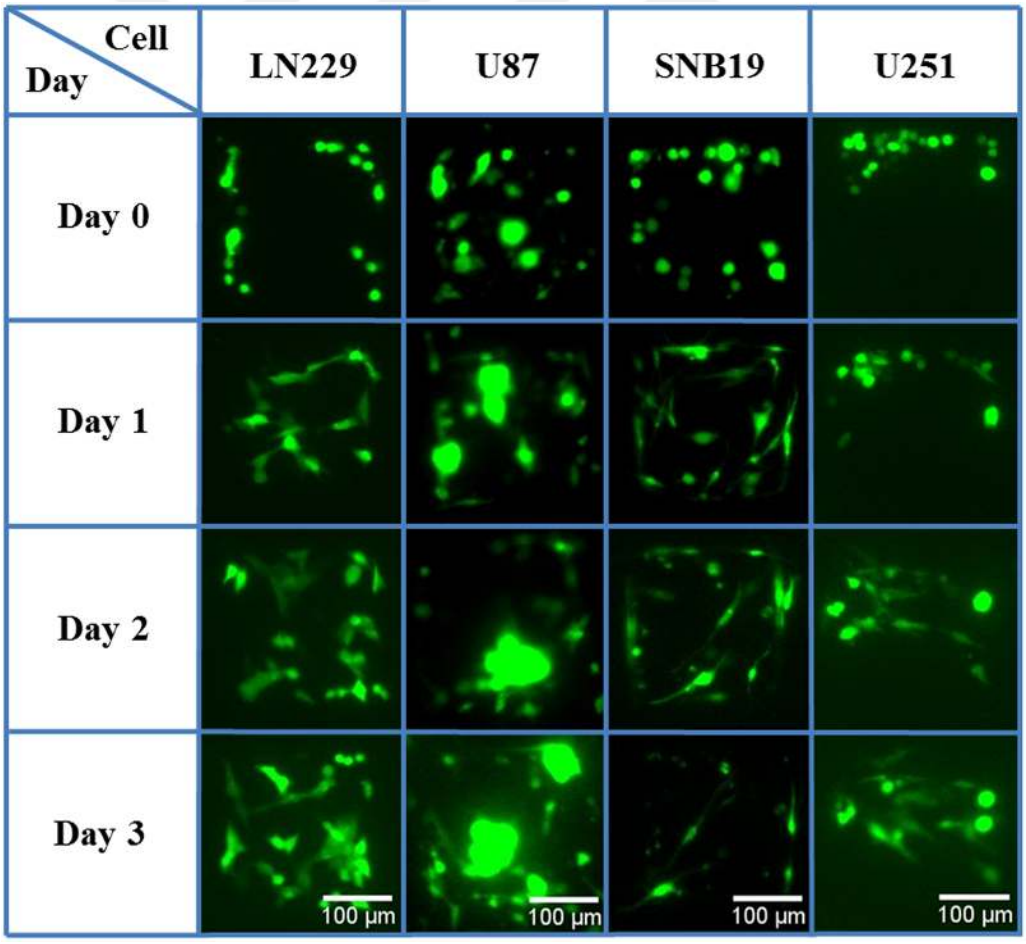

Figure 1.6: Evolution of cell shape of GB Cells in *in Vitro* ECM (Diao et al., 2019)

1.15 2D vs 3D cell culture

The traditional 2D *in vitro* cell proliferation method (on a flat substrate) has been employed in cancer research for an extended period. Nonetheless, this approach cannot be entirely transposed into clinical studies to accurately reflect physiological conditions. This culture fails to replicate the authentic TME owing to insufficient cellular communication and interaction (Habanjar et al., 2021). To address these limitations, 3D culture systems are being increasingly developed in research and have become essential for tumor research, tissue engineering, and fundamental biological studies. 3D culture has attracted considerable interest in biomedicine due to its ability to emulate tissue architecture and functions. The 3D matrix offers a dynamic structure in which its elements are deposited, destroyed, or altered to define functions and serve as a substrate for cell attachment, facilitating processes such as adhesion, proliferation, communication, and apoptosis.

In 3D environments, cells reorganize microspheres to form aggregates, highlighting the advantages of microspheres in such settings and enabling cells to adapt the environment to suit their needs (Clara-Trujillo et al., 2019). The following table (Table 1.3) presents a comparative overview of 2D and 3D *in vitro* cell culture techniques in cancer research, highlighting key biological and functional discrepancies between the two systems.

Table 1.3: The main distinctions between 2D and 3D *in vitro* cell culture techniques (Louit, Galbraith, & Berthod, 2023)

Distinctions between 3D and 2D	
3D	**2D**
Resemblance to levels of gene and protein expression in vivo	Activation of apoptosis
Improved maintenance and differentiation of cells	Hypersensitivity to drugs
Precise reaction to two-dimensional mechanical stimuli	Homogeneous distribution of the culture media
Different combinations of cell types	Excessive growth
Maintained its natural form	

Traditional 2D technologies are inadequate for accurately mimicking real tissue and representing disease pathophysiology. Conversely, 3D cell culture models provide a novel approach for researchers by creating tissue-like 3D environments that enhance cellular performance and survivability (Keles et al., 2025).

1.16 Cell interactions and behaviors

Cells are three-dimensionally encased by the ECM within intricate *in vivo* microenvironments, engaging with both the ECM and adjacent cells. Therefore, replicating the ECM environment is crucial for efficient cell culture models. A variety of natural and synthetic hydrogels have been utilized to mimic ECM settings, considering their physical, chemical, and biological properties, such as biocompatibility, biodegradability, and biochemical functional groups. Due to these characteristics, hydrogels have been integrated with micro technologies and utilized in organ-on-a-chip applications to more precisely emulate the *in vivo* microenvironment (S. H. Lee, Shim, Kim, & Sung, 2017). Assessing the migratory and invasive potential of tumor and stromal cells is crucial for advancing cancer diagnosis, prognosis, medication development, and treatment. *In vitro* migration and invasion assays encompass a diverse range of methodologies for studying cell migration and invasion, providing a concise overview of recognized assays, along with their benefits, drawbacks, and a visual representation of their underlying concepts. It advocates adopting exclusion zone tests in conventional applications involving adherent cells, emphasizing its advantages, including straightforward setup and simplicity of analysis. Furthermore, it examines advanced 3D migration and invasion models that accurately replicate *in vivo* settings, potentially uncovering novel pathways underlying the dissemination of cancer cells. The subsequent figure (Figure 1.7) depicts the predominant *in vitro* migration assay models that facilitate the methodical examination of cell motility under regulated conditions (Kramer et al., 2013). These models encompass the Boyden chamber (Transwell) assay, a widely used method for measuring directed cell migration through a porous membrane, and the wound healing assay, which evaluates collective migration in a 2D setting. Other techniques, such as the cell exclusion zone assay and the fence (or ring) assay, create spatial limits that facilitate the study of cell movement. The microcarrier bead assay and the spheroid migration assay are two specialized methods that allow cells to move away from 3D structures. Capillary-based techniques, such as the capillary chamber migration assay, horizontal capillary assay, and capillary tube migration assay, provide accurate microenvironmental regulation. Conversely, the leukocyte migration agarose assay is especially effective for examining immune cell motility in semi-solid matrices.

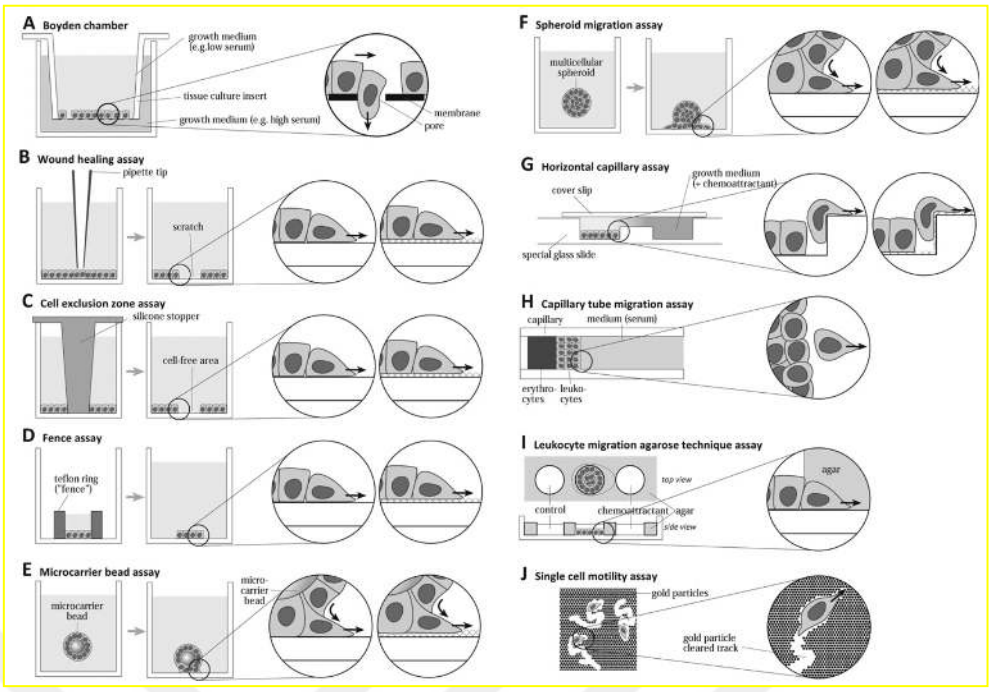

Figure 1.7: Commonly used *In Vitro* migration assays (Kramer et al., 2013)

On the other hand, the following figure (Figure 1.8) shows the commonly used *in vitro* invasion experimental models, which are listed as follows (Kramer et al., 2013).

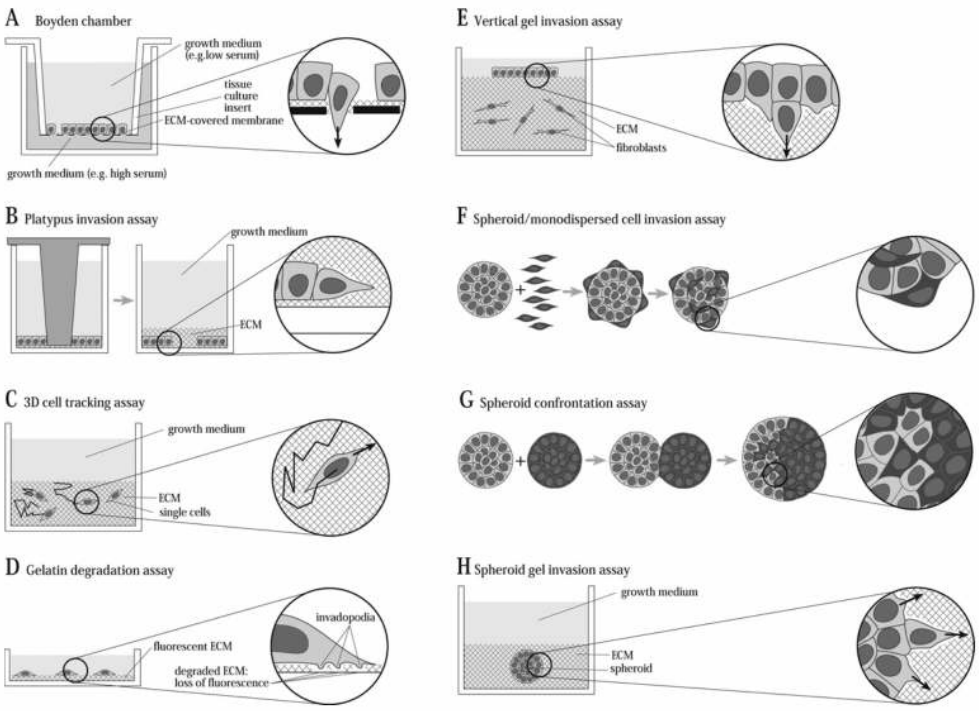

Figure 1.8: Commonly used *In Vitro* invasion assays (Kramer et al., 2013)

The following table (Table 1.4) combines the results of *in vitro* experimental studies for both migration and invasion assays.

Table 1.4: Commonly used migration and invasion assays *In Vitro*

In Vitro Experimental Set-Ups	
Migration Assays	Invasion Assays
Boyden chamber assay (Transwell)	Transwell invasion assay
Wound healing assay	Platypus invasion assay
Cell execution zone assay	3D cell tracking
Fence essay (ring)	Gelatin degradation assay
Micro carrier bead assay	Vertical gel 3D migration/invasion assays
Spheroid migration assay	Spheroid/monodispersed cell invasion assay
Capillary chamber migration assay	Spheroid confrontation assay
Horizontal capillary assay	Spheroid gel invasion assays
Capillary tube migration assay	
Leukocyte migration agarose assay	

1.17 Disease models

Enhancing cancer therapy requires the use of biomimetic 3D *in vitro* studies, such as tumoroids and spheroids (Seyfoori, Liu, Caruncho, Walter, & Akbari, 2025). The TME dictates disease advancement and affects treatment efficacy. Many aggressive solid tumors demonstrate increased infiltration of myeloid cells, particularly tumor-associated macrophages (TAM). An extensive examination of the connections between many biological elements of the tumor microenvironment and the ECM is crucial for clarifying the processes of disease advancement. This approach is particularly crucial for predicting therapeutic responses to conventional treatments and evaluating the effectiveness of TME-targeting pharmaceuticals (Rebelo et al., 2018).

1.17.1 Cancer

The TME is crucial to tumor start, development, and responses to treatment. A significant portion of our knowledge regarding tumors and their microenvironments is derived from diverse cell culture techniques. Biomimetic 3D culture systems, which replicate the *in vivo* microenvironment, are poised to serve as precise tools for diagnosis, drug evaluation, and personalized medicine (Atat et al., 2022). The following

figure (Figure 1.9) demonstrates an overview of 3D cancer cell culture systems and limitations of conventional 2D cultures. Cancer cells are sustained and cultivated in 3D culture systems utilizing various techniques, including scaffold-based approaches, bioreactors, organoids, microfluidic devices, and tumor spheroids, which are cellular aggregates that mimic the *in vivo* development of tumor cells. Preclinical drug testing and illness modeling are two current applications for 3D systems.

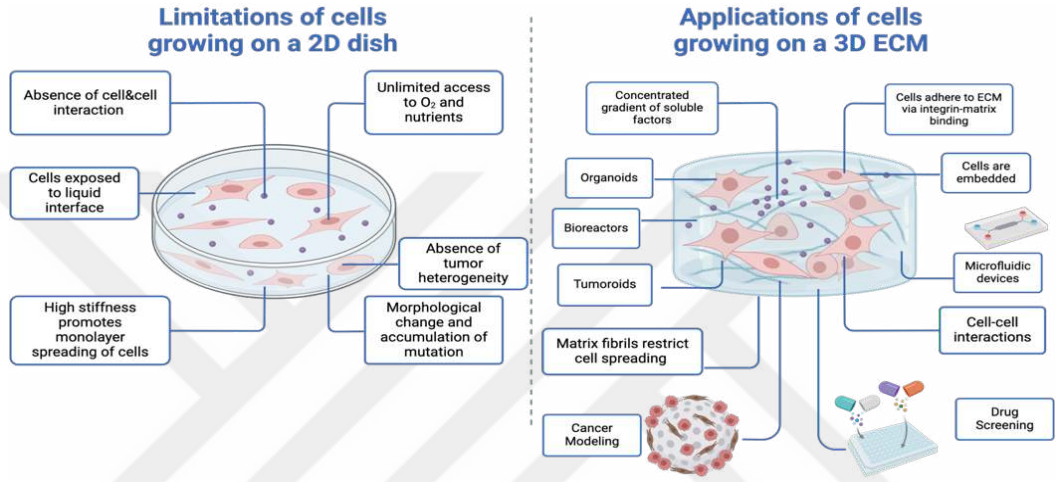

Figure 1.9: Overview of 3D cell culture systems and limitations of conventional 2D cultures in cancer research

Cell culture is crucial for pharmaceutical discovery, oncology research, and stem cell studies. 3D cell culture techniques facilitate more sophisticated research, yielding precise data on cellular connections, tumor attributes, and disease mechanisms. Scaffold-based techniques, including hydrogel and polymeric solid supports, as well as scaffold-free technologies such as hanging drop microplates and magnetic levitation, offer unique benefits and applications. 3D cell culture may serve as a conduit between 2D and animal models, offering particular methodologies for investigating organ functionality (Jensen & Teng, 2020).

1.17.2 Glaucoma

Glaucoma represents a primary global reason for irreversible vision impairment (Dumbrăveanu, Cușnir, & Bobescu, 2021). The relative afferent pupillary deficit (RAPD) diagnosis, as determined by the swinging flashlight test with neutral density

filters, demonstrated acceptable sensitivity and high specificity for glaucomatous optic neuropathy; thus, the swinging flashlight test is a potentially valuable screening tool (Charalel, Lin, & Singh, 2014). Similar tests may be investigated using wearable products, such as headsets or smart glasses, which simultaneously capture eye movements and compare the consistency of pupil movements. A software program may act as a pupilometer to detect pupil radius.

3D cell culture platforms facilitate precise and dependable modeling of tissue-like interactions between cells and the ECM. They integrate conventional 2D techniques with *in vivo* and *ex vivo* models, potentially aiding in the elucidation of causal processes underlying outflow challenges in ocular hypertension and glaucoma. The Matrigel 3D culture platform may facilitate the development and screening of glaucoma drugs (Ghosh & Herberg, 2024).

DL-based glaucoma detection using CNN and digital fundus images may present a promising method for accurate diagnosis; however, such AI investigations require substantial improvements to enhance the clarity, coherence, and scientific rigor of the results (Song, Wang, & Xing, 2024).

1.17.3 Neurodegenerative diseases

Alzheimer's disease (AD), Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), and Huntington's disease (HD) are the most prevalent neurodegenerative illnesses. In these diseases, glial cells can proliferate rapidly, and cells can be extracted from an embryonic or postnatal brain. Neurodegenerative conditions can be simulated using scaffolds, spheroids, organoids, 3D bioprinting, microfluidic systems, and organ-on-a-chip methodologies (Louit et al., 2023).

1.17.4 Alzheimer’s disease

Amyloid-β plaques and neurofibrillary tangles are two histological indicators of Alzheimer's disease (AD), the most common form of dementia (Mancino et al., 2018). Amyloid-β plaques and neurofibrillary tangles were not present in human neurons isolated from AD patients; nonetheless, there was an elevation in phosphorylated tau

and toxic amyloid-β species. We report that in a 3D culture system derived from human brain stem cells, familial AD mutations in β-amyloid precursor protein and presenilin-1 can induce significant extracellular deposition of amyloid-β, including the formation of amyloid-β plaques. This distinctive technique may facilitate the development of more accurate human neural cell models for various neurodegenerative diseases (S. H. Choi et al., 2014).

1.17.5 Parkinson's disease

Parkinson's disease (PD) is the second most prevalent neurodegenerative disorder, significantly diminishing the quality of life for sufferers due to its extensive array of motor and non-motor symptoms (Keles et al., 2025). Substantial loss of dopaminergic neurons and the aggregation of α-synuclein protein into ubiquitin-positive Lewy bodies represent a fundamental neuropathological feature of PD. Scientists are investigating the prediction of PD subtypes based on AI algorithms (D'Sa et al., 2023). The following figure (Figure 1.10) outlines the Workflow for machine learning with image data using a CNN to predict subtypes of PD. Researchers have designed and trained a dense neural network using Python and the TensorFlow framework (Figure 1.10). The model's architecture had three fully connected layers with ReLU activation functions. Each layer was followed by a dropout layer with a dropout rate of 0.2, which helped prevent overfitting. The last layer of the network was a dense layer that utilized a SoftMax activation function to enable multiclass classification.

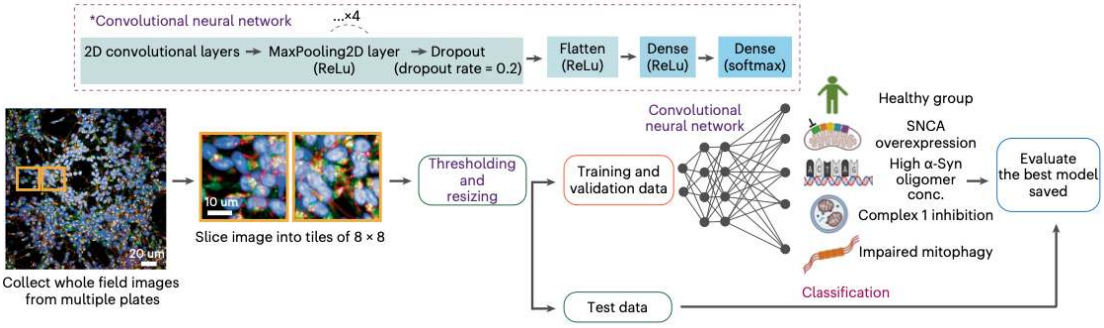

Figure 1.10: Workflow for machine learning with image data via CNN for predicting subtypes of PD (D'Sa et al., 2023)

The generation of significant nuclear-associated Lewy bodies from endogenous wild-type α-synuclein, which is translationally regulated by its own promoter in human cell culture models, requires costly and labor-intensive techniques. Scientists have shown that fully developed human SH-SY5Y neuroblastoma cells can grow in biomimetic 3D ECM platforms with Lewy-body-like pathology when exposed to exogenous α-synuclein species (Taylor-Whiteley, Le Maitre, Duce, Dalton, & Smith, 2019).

1.17.6 ALS

A human neuromuscular co-culture system was established to investigate the functionality and maturation of human motoneurons (MNs) and their corresponding muscle cells at neuromuscular junctions (NMJs). The system, composed of induced pluripotent stem cell (iPSC)-derived motor neurons and 3D skeletal muscle tissue, was utilized as an *in vitro* model to investigate the pathophysiology of amyotrophic lateral sclerosis (ALS). The system exhibited diminished neuromuscular connection and muscle contraction in co-cultures, including motor neurons with the ALS-associated SOD1 mutation (Massih et al., 2023).

A separate study utilized iPSC-derived cells to model ALS and investigate the previously unreported phenotype. Wild-type iPSC lines were created, and C9orf72 knockdown human iPSCs were differentiated into astrocytes, motor neurons, and spinal cord organoids (SCOs). The C9-SCOs exhibited characteristics of spinal cord development and various neuronal cell types. Proinflammatory factors were elevated in C9-iPSC-derived 2D cells and 3D cultivated SCOs, indicating reduced C9orf72 expression and pathological features of ALS (Guo et al., 2024).

1.17.7 Pancreatic adenocarcinoma

Co-culturing pancreatic adenocarcinoma (PDAC) tumoroids with human-derived fibroblasts demonstrates the pro-invasive effects of stromal components on tumor growth and drug response. This investigation highlights the significance of advanced biomimetic 3D models, such as tumoroids, in enhancing the comprehension of cancer complexity and therapeutic responses (Seyfoori et al., 2025).

1.17.8 Glioblastoma

Recently, 3D cultures have provided a means to address these challenges, and brain organoids are emerging as a groundbreaking tool in glioblastoma research. The capacity to effectively generate brain organoids through gene editing and to conduct co-cultures with PDT spheroids has facilitated the exploration of cancer progression in a manner that more accurately reflects brain tissue architecture. Moreover, the establishment of biobanks with organoids derived from GB patients is essential for advancing precision medicine therapy (Andreatta et al., 2020).

GB TME is notably heterogeneous, comprising a mixture of malignant and non-cancerous cells, with non-cellular components such as the ECM. The following figure (Figure 1.11) illustrates continuous, multi-directional interactions between glioblastoma cells and normal stromal cells, including endothelial cells and astrocytes, which enable the tumor to exploit its surrounding microenvironment and induce a tumor-supportive phenotype. Extracellular vesicles have lately been identified as directly participating in this interaction, alongside diffusible cytokines and growth factors (Simon, Jackson, & Giamas, 2020).

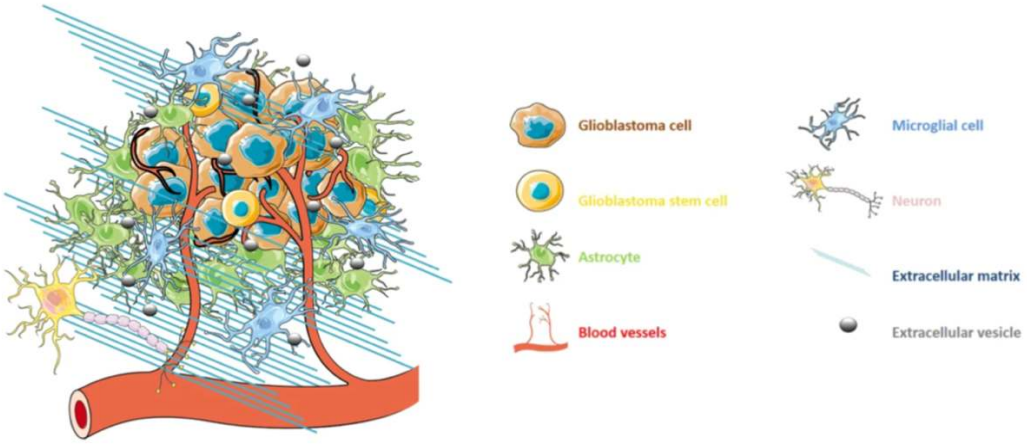

Figure 1.11: The heterogeneous nature of GB cells in their micro-environment (Simon et al., 2020)

The advanced tumoroid model facilitates the evaluation of various GB ECM compositions, including collagen and Reelin, to determine their impact on GB

invasiveness. The tumoroid *in vitro* model facilitates the screening of chemotherapeutics, including TMZ with iron chelators, in both monotherapy and combination therapy situations. This approach assesses their toxicity on the viability and apoptosis of the GB model embedded within a complex ECM (Seyfoori et al., 2025).

The TME plays an essential role in tumor development, progression, and responses to chemotherapy. A significant portion of our knowledge regarding tumors and their microenvironments is derived from diverse cell culture techniques. 3D culture systems, which replicate the *in vivo* microenvironment, are poised to serve as precise tools for diagnosis, drug evaluation, and personalized medicine (Ratliff et al., 2023).

1.18 Applications of AI in GB

Although this thesis examines the integration of 3D technologies and AI in analyzing the GB TME, we have explored the status of AI applications in overall GB research in this chapter. AI applications applied to GB include diagnosis, treatment, determining severity, progression, and categorization (Luo et al., 2023).

AI is essential in many aspects of GBs, including tumor categorization, diagnosis, differentiation, scoring, treatment response, clinical outcome forecasting (prognosis and recurrence), molecular characteristics, clinical categorization, TME characterization, and drug development (Figure 1.12). The implementation of AI in GB has progressed significantly, especially in medical image analysis; however, AI is still in its initial phase of validation and deployment (Luo et al., 2023).

Some studies focus on AI within 3D cell culture technologies, aiming to enhance our understanding of the TME, which could facilitate the development of customized treatments and personalized medicine (Walsh & Quail, 2023). The various cellular and acellular components constituting the TME may collectively influence tumor growth, invasion, metastasis, and treatment responsiveness (Elhanani et al., 2023). It is becoming more crucial to observe the TME's cellular diversity and 3D structure to understand how diseases progress and how treatments work. The following

representation (Figure 1.12) outlines the utilization of AI in numerous aspects of GB, including tumor classification, diagnosis, differentiation, scoring, effectiveness of therapy, clinical outcome prediction (prognosis and recurrence), molecular characteristics, clinical classification, TME analysis, and drug development (Luo et al., 2023).

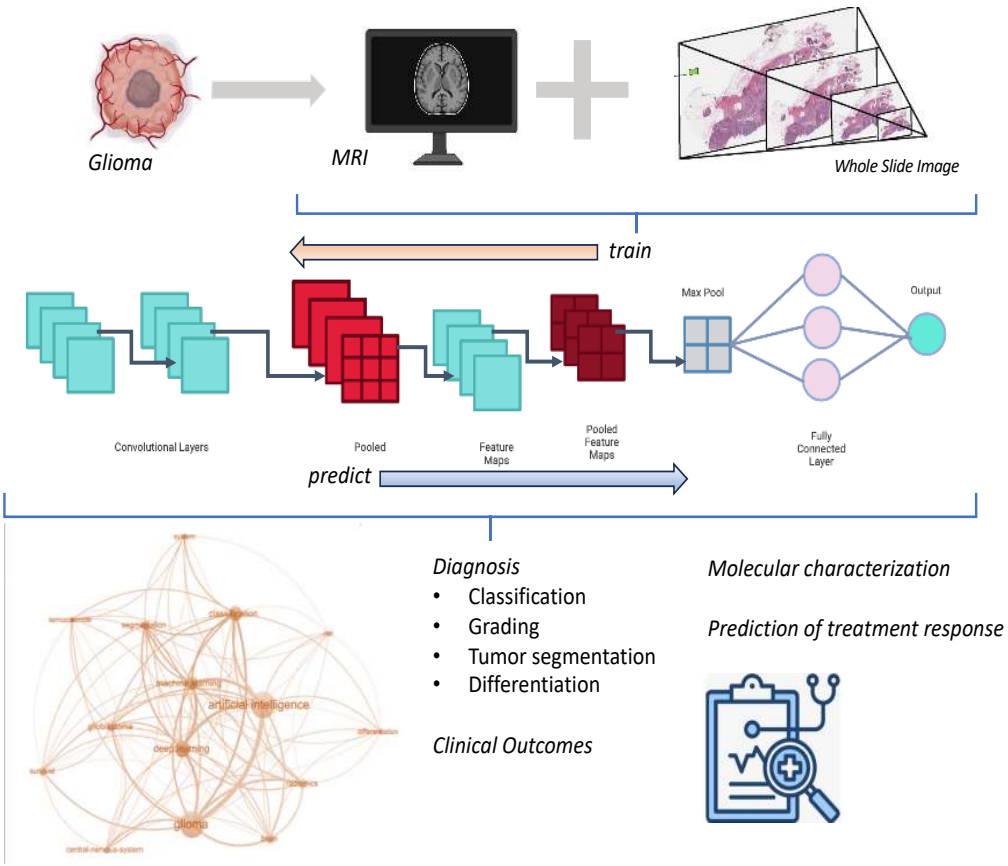

Figure 1.12: Utilization of AI in GB (Luo et al., 2023)

The two methods presented here offer distinct approaches to drug testing and brain tumor research (Luo et al., 2023). Studies are offering a comprehensive survey of AI approaches employed in brain tumor images, including a narrative investigation that provides a general summary of these techniques. Some studies explain an experimental approach that focuses on personalized drug testing using machine learning, while others provide an overarching view of AI methodologies employed in brain tumor imaging.

Gliomas, the most prevalent malignant neoplasms in the central nervous system, present considerable challenges in diagnosis and management. MRI has advanced considerably with the introduction of multimodal MRI, providing a comprehensive perspective on tumor intricacies. This comprehensive approach enhances patient outcomes by providing accurate diagnoses and customized treatment plans. The integration of artificial intelligence with MRI data is expected to revolutionize the field, enabling more sophisticated studies (Y. Yan et al., 2025).

Glioblastomas represent the predominant kind of malignant brain tumors, resulting in around 200,000 fatalities each year. The present therapeutic modalities comprise surgical excision, radiation therapy, and chemotherapeutic agents. Precise tumor segmentation is essential for efficient treatment planning. Nevertheless, existing methodologies are laborious, subjective, and costly. A diagnostic tool powered by AI has been developed to autonomously distinguish between brain tumors and predict the promoter methylation status using brain MRI data. This approach enables physicians to submit brain MRI images, provide segmentation and methylation predictions, and guide treatment planning, ultimately improving survival time (Goddla, 2022).

The studies based on AI and GB demonstrate how combining 3D *in vitro* technology with AI and multi-omics datasets can revolutionize our knowledge of the TME. This kind of integration could transform personalized cancer treatment and clinical decision-making. However, studies emphasize the importance of optimizing these technologies for clinical applications and increasing researcher accessibility. Even while AI has much potential to help with GB patients' clinical care and precision medicine, there are still several obstacles to overcome, especially regarding deployment. Some studies present a novel method that enhances the capacity to evaluate GB treatment responses and microenvironment features by integrating 3D bioprinting with a multi-algorithm machine learning technique. This combined strategy may be effective for a wide range of cancer types. These studies also highlight how AI is becoming increasingly valuable in brain tumor imaging, enabling the planning of treatments with higher precision and the prediction of patient outcomes with superior accuracy. The goal of this investigation is personalized medicine (Scimeca, Urbano, Toschi, Bonanno, & Schillaci, 2021).

1.19 Application of AI in 3D *in vitro* technologies on GB TME

Organoid intelligence in biocomputing could yield valuable insights into neuroscience while presenting a biohybrid approach to information processing. Progress in the engineering of brain region-specific organoids, the development of sensors and signal-processing tools, and the integration of artificial intelligence, along with improvements in downsizing, will enable the development of organoid intelligence, impacting biomedicine and other domains (Smirnova, 2024). This section provides an overview of the methodologies utilized to investigate the applications of AI in studying the GB TME using 3D technologies. The TME is a dynamic and complex environment, containing an extensive range of cell types and extracellular materials. Comprehending the TME's cell-level heterogeneity and spatial arrangement is essential for elucidating the pathways that drive tumor growth and treatment responses (Walsh & Quail, 2023).

The brain consists of several cell types that sustain the integrity of normal brain function through their meticulous coordination. The interaction of GB cells with several cell types (immune cells, neurons, astrocytes, pericytes, endothelial cells, etc.) establishes a distinctive milieu crucial for GB growth and invasion (L. L. Wang et al., 2021). Understanding the architecture and evolution of the TME may improve the prevention and treatment of GB. AI offers a technique for a noninvasive sampling of the tumor microenvironment, facilitating the dynamic and comprehensive assessment of geographically diverse glioblastoma. Recent research has demonstrated that an ANN model employing DNA damage response genes can predict outcomes in LGG, exhibiting significant robustness and superiority in response to mutation and immunological microenvironment conditions (Luo et al., 2023).

The primary objective of GB research is to identify cellular subpopulations within the TME associated with prolonged clinical outcomes, intending to develop novel medicines and diagnostics (Cè et al., 2023). Deep CNNs for segmenting whole-slide pathological images identify the interactions of microglia, pericytes, and monocytes with tumor cells in regions of tumor and cellular malignancy, as well as microvascular growth, which correlates with adverse prognostic outcomes in GB.

Several 3D model systems, including spheroids, organoids, and bio-printed scaffolds, were employed to mimic the intricate nature of the GB TME (Piantino, Figarol, & Matsusaki, 2021). The 3D systems were chosen for their capacity to replicate relevant aspects of the GB microenvironment, including tumor heterogeneity, cellular interactions, and an ECM (Bhattacharya et al., 2024). The choice of the experimental model was based on the research's specific goals. Spheroids were utilized for high-throughput screening because they are easily formed, quickly develop, and are well-suited for large-scale tests. On the other hand, organoids had a more sophisticated and physiologically appropriate multicellular structure, which enabled the detailed study of tumor-stroma dynamics and interactions with immune cells.

Additionally, 3D bio-printed structures were incorporated to create bespoke microenvironments with a regulated spatial organization of various cell types. This approach facilitated the representation of tissue structure and functional heterogeneity. Glioblastoma cells and normal stromal cells, including endothelial cells or astrocytes, talk to one another all the time in many directions. Such communication enables the tumor to take over its microenvironment, resulting in a tumor-supportive phenotype. Recent descriptions directly implicate extracellular vesicles in this interaction, along with diffusible cytokines and growth hormones.

Obtaining Data from 3D Models is enhanced by imaging and monitoring methods. Advanced imaging techniques are used to collect real-time data from the 3D cultures (Alieva et al., 2024). Researchers have used advanced imaging techniques to provide a clear picture of how cells are organized and behave in 3D biological systems. Confocal microscopy was utilized to produce optically sectioned pictures that selectively capture in-focus planes, facilitating a profound analysis of cellular structures and their interactions while eliminating out-of-focus light through the application of one or more apertures (Nikon, 2024). In addition to this method, live-cell imaging has enabled the real-time observation of dynamic processes, such as cell proliferation, invasive behavior, and treatment responses. Fluorescent markers were also used to label and keep track of specific groups of cells, including cancer stem cells and immunological subsets. This enabled the accurate mapping of cellular dynamics in space and time within complex biomimetic 3D models.

Deep learning models are employed to leverage the optimal features of confocal microscopy. Scientists train the U-Net and ResU-Net convolutional neural network (CNN) architectures to anticipate the appearance of high-quality photographs. In the following figure (Figure 1.13), a schematic representation of the confocal fluorescence microscopy is presented (Suresh et al., 2025). This diagram illustrates the operation of a confocal microscope system's optics. This system can take high-resolution pictures in three dimensions (`x, y`, and `z`). The laser beam enters the system and is directed to the primary lens, where it is precisely focused. After passing through the lens, the beam hits a beam splitter, which serves two functions: it sends some of the light down the detecting pathway and reflects the rest downward toward the objective lens. The objective lens then focuses the laser into the sample, which is placed directly below it to ensure that the intended focal volume is excited appropriately. The light that comes from the sample or is reflected off it goes back through the objective and beam splitter before being projected onto the detection screen, écran. This setup enables the system to take optically sectioned pictures at particular depths, allowing for the creation of detailed 3D datasets by scanning along the x-, y-, and z-axes in a controlled manner.

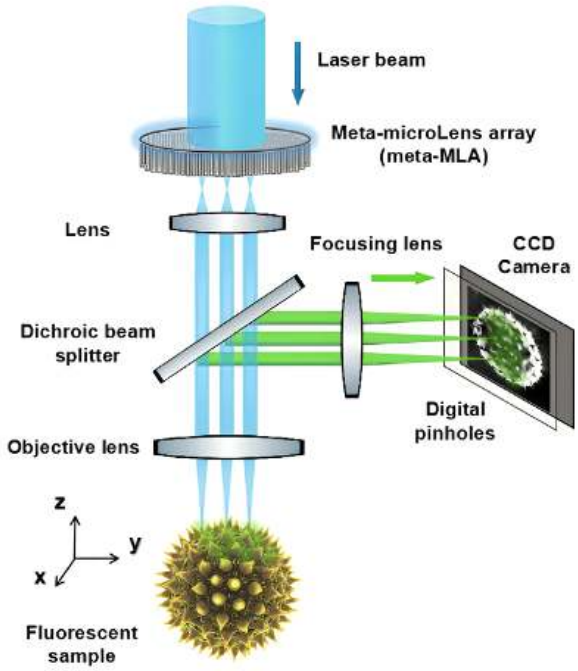

Figure 1.13: Schematic diagram of the meta-MLA confocal microscopy (Suresh et al., 2025)

The AI integration with 3D technologies encompasses various stages, from developing biologically significant 3D tumor cultures and progressing to utilizing AI algorithms for data analysis, prediction, and therapeutic exploration (Boehm et al., 2022).

1.20 Patient-derived tumor cultures & machine learning workflow (GlioML)

In literature, studies have been conducted to obtain patient-derived tumor (PDT) samples from GB patients, which are then used to create PDT cultures for drug testing. The cultures are bio-printed using a custom bio-ink made of gelatin methacrylate (GelMA) and hyaluronic acid methacrylate (HAMA) to recreate the GB ECM more precisely. The bio-ink, composed of a chemical HAMA and GelMA, has been designed to replicate the GB ECM commonly used in organoid 3D technologies. Scientists have developed a multi-algorithm machine learning workflow called GlioML to predict drug responses. GlioML employs nine supervised learning algorithms and a weighted ensemble model, thereby adopting a data-driven, experimental approach for personalized drug testing (Min Tang et al., 2024).

GB cells were acquired from either PDT tissues or established commercial GB cell lines, and subsequently proliferated under specific culture conditions to create 3D model systems. To generate spheroids, GB cells were placed in non-adherent culture plates, where they could not adhere to any other surface, allowing them to form tight, multicellular spheroids naturally. Using stem cell-derived GB structures implanted in Matrigel, organoid models were made to mimic substantial parts of the native ECM and facilitate the intricate assembly of tissues. Simultaneously, 3D bio-printed GB models were created by accurately placing GB cells, stromal components, and matrix materials utilizing a controlled printing method. After being printed, these structures were transferred to bioreactors designed to maintain optimal environmental conditions, which helped the cells remain alive, the structures remain strong, and the physiological relevance across the various 3D platforms.

The subsequent graphic (Figure 1.14) depicts the research examining the feasibility of using 3D bio-printed models with GlioML to forecast and evaluate personalized therapy responses in glioblastoma patients. The graphic depicts a comprehensive methodology that integrates an experimental component involving PDT creation and the assessment

of various compounds and drugs with a computational facet (collection of genetic expression data and GlioML prediction) (Min Tang et al., 2024).

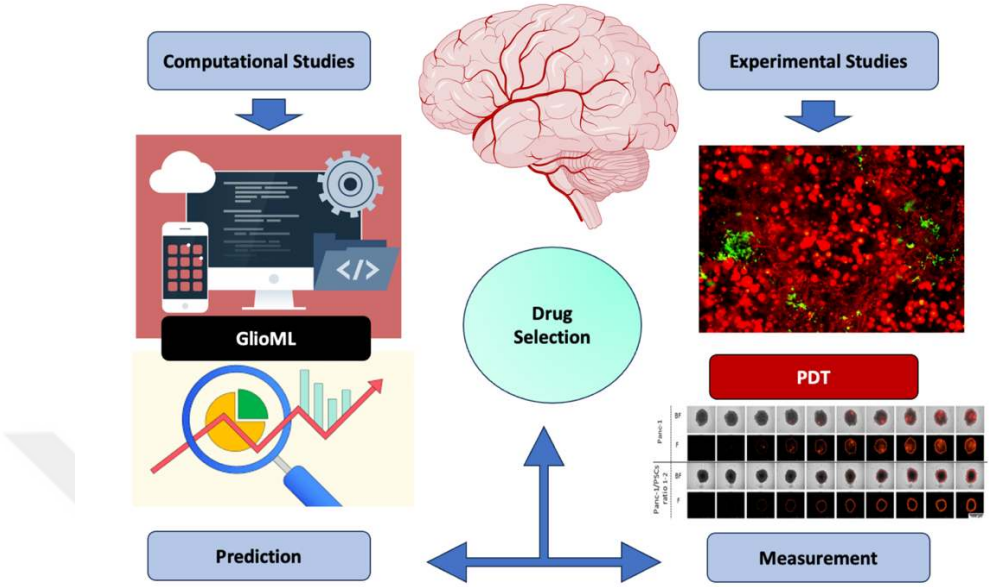

Figure 1.14: Combination of Experimental and Computational Studies. The illustration presents the research that explores the integration of 3D bio-printed models with GlioML to predict and evaluate individual treatment responses in GB patients (Min Tang et al., 2024)

The following graph (Figure 1.15) represents GlioML's PCA of the area under the curve (AUC) matrix differentiated between the WHO score GB III and GB IV, revealing particular medicament treatment response patterns. Three GlioML drugs, dasatinib, RSL, and lovastatin, showed positive responses in CW468 cells, along with median viabilities exceeding those with TMZ. These compounds were found in three clusters. The t-SNE map indicates that dasatinib and RSL cluster together, while lovastatin and TMZ occupy distinct clusters based on the responses of the 481 compounds to patient samples (Min Tang et al., 2024).

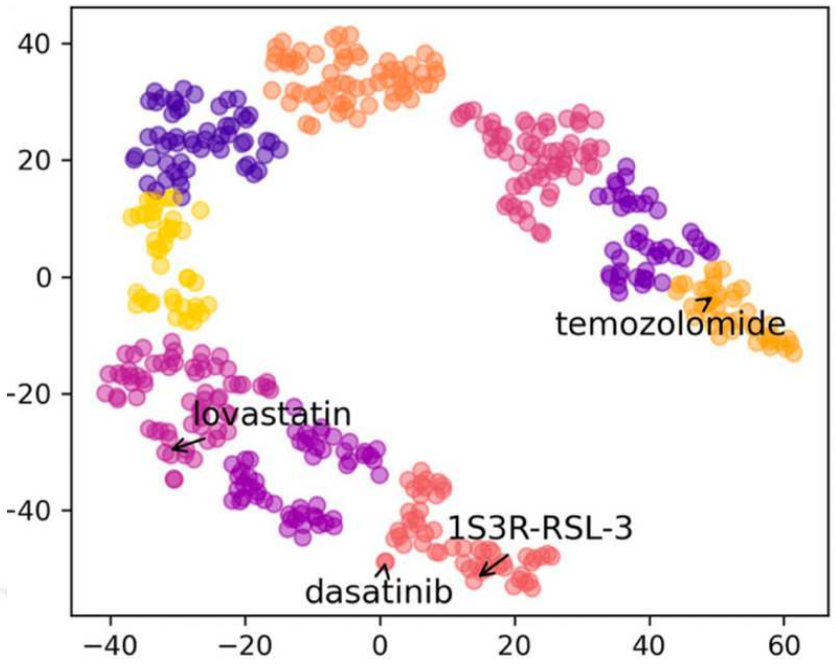

Figure 1.15: Drug Responses of GB Patient Samples (Min Tang et al., 2024)

The following figure (Figure 1.16) illustrates the raw data on the expression of protein-coding genes, which have been transformed into single-sample analysis of gene set enrichment (ssGSEA) results for the selected gene sets. In contrast with the gene expression results, the ssGSEA scores offered enhanced interpretability and elucidated the regulatory extent within each gene set. Sixty-six cell lines from the CTRP were thoroughly identified within the CCLE and subsequently isolated for further utilization. The AUC for each cell-compound pairing functioned as labels, while unlabeled couples were pre-processed and subsequently excluded. We consolidated all this information to produce the preliminary training data for the workflow (Min Tang et al., 2024).

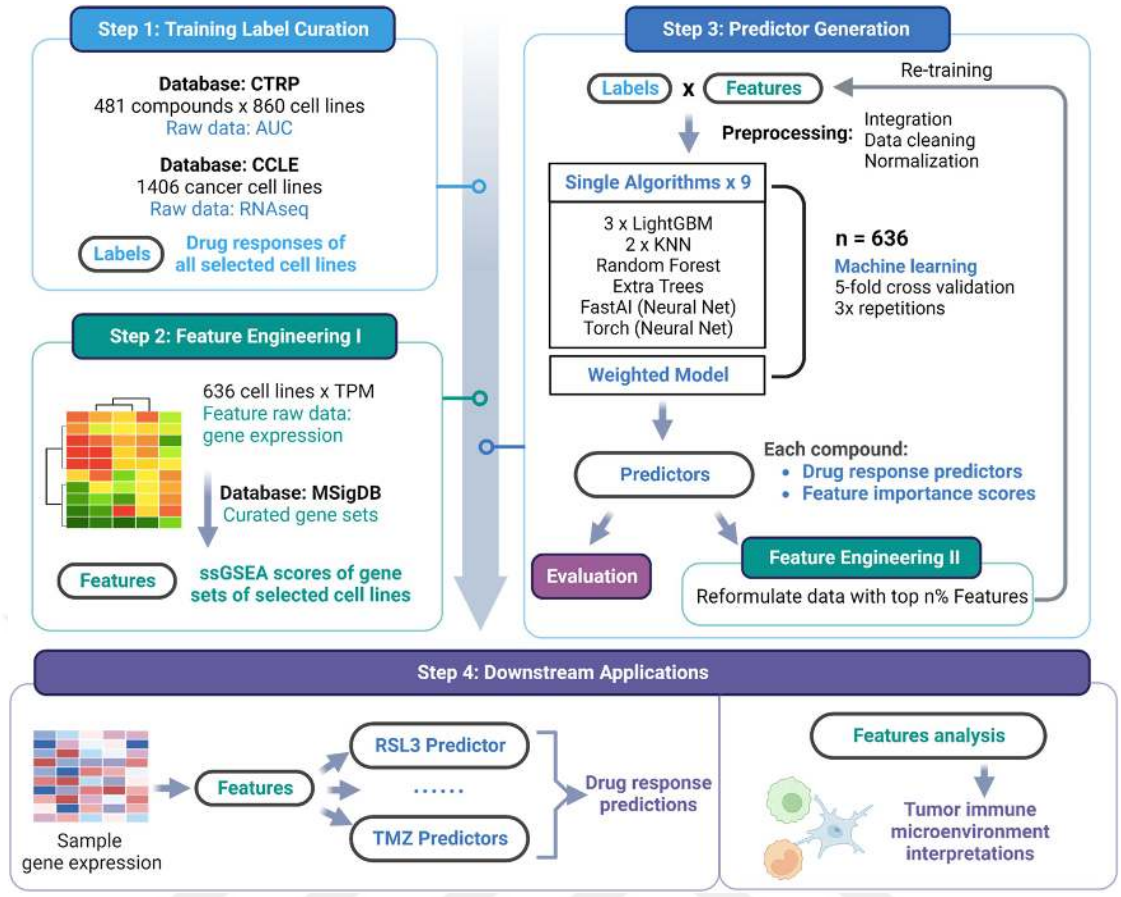

Figure 1.16: A machine-learning technique to predict medication responses, GlioML. This illustration represents the GlioML algorithm, an ML approach that utilizes multiple algorithms to predict pharmaceutical responses. The process comprises four primary elements (1) Data preprocessing, which includes labeling and standardizing a training dataset; (2) Feature engineering, the discovery of pertinent variables for precise prediction of GB treatment response; Training Module: an extensive procedure for developing individual predictors, involving feature reduction, initial training, and final performance assessment; lastly (4) Downstream Applications, which indicate features involve microenvironment analysis and treatment response prediction (Min Tang et al., 2024)

1.21 Multi-omics technologies for elucidating tumor immune dynamics

Single-cell proteomics, epigenomics, transcriptomics, and metabolomics are single-cell spatial omics methodologies designed to evaluate tumor slices and tissues in two dimensions. The following graph (Figure 1.17) illustrates various data layers that can be

combined into a single sample. Emerging technologies are being developed that enhance tissue-clearing processes or serial segment alignment, broadening these capabilities into three dimensions to encompass whole tissue volumes or entire organisms (Walsh & Quail, 2023).

Additionally, time-controlled barcoding or sampling methods incorporate spatiotemporal evaluation of tumor immune system responses across time at a single-cell level within these datasets for data enrichment (Z. Xie et al., 2021). The objective is to obtain omics data from the 3D GB models, which were collected using high-throughput technologies, to create a comprehensive molecular dataset. To obtain a more comprehensive characterization of glioblastoma (GB) biology and its associated TME, several analytical methodologies must be further refined and integrated (Wilkinson et al., 2016).

Gene expression profiling using RNA sequencing facilitates an in-depth analysis of transcriptional programs within both tumor and stromal compartments, elucidating the molecular determinants of disease progression (Sammut et al., 2022). In addition to transcriptomic analysis, proteomic methods can measure the amount of protein present and investigate substantial signaling pathways that are active in GB tissues. This gives genomic and transcriptomic results a functional context (L. Chen, Qin, Guo, Wang, & Li, 2021). Single-cell sequencing is also a significant step forward in understanding the diverse range of cell types that comprise the TME. It enables the accurate identification of separate subpopulations and their functional characteristics (J. H. Lee et al., 2015). In general, improving these methods will significantly enhance molecular resolution and help us gain a deeper understanding of how GB works. Examples of technologies within each category are shown in the figure below (Figure 1.17) (Walsh & Quail, 2023).

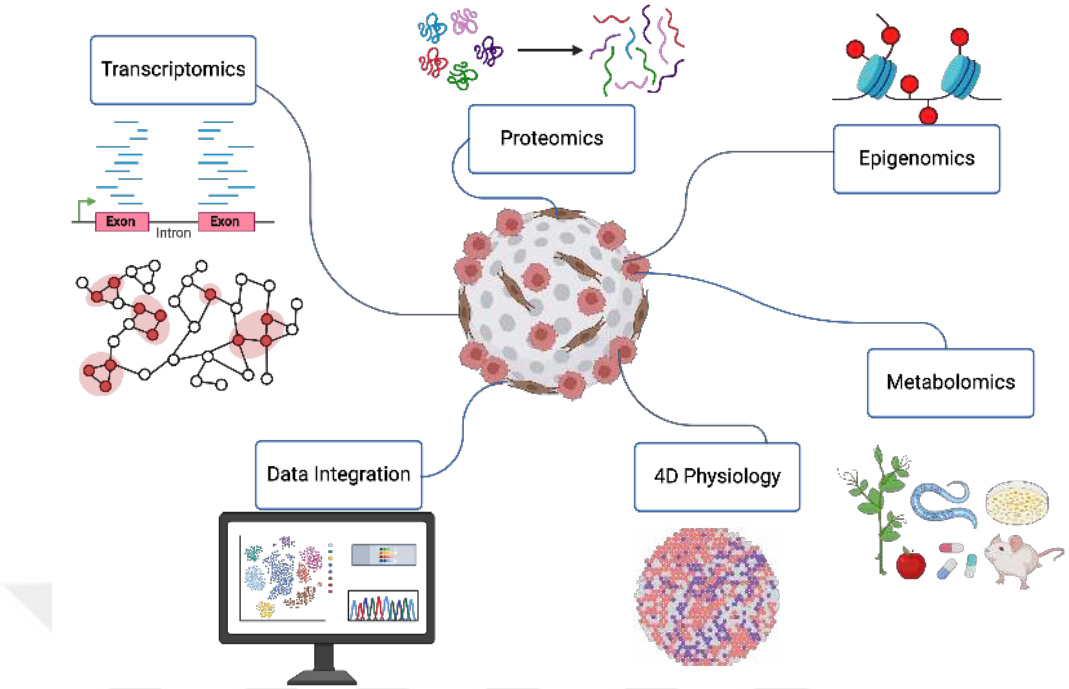

Figure 1.17: Omics Technologies (Madonna et al., 2022. , Walsh & Quail, 2023)

1.22 Next-generation sequencing technology

Next-generation sequencing (NGS) tools have significantly enhanced our comprehension of the GB TME (J. H. Lee et al., 2014). Scientists have employed machine learning techniques to illustrate the predictive significance of monocytes in the TME and their correlation with survival outcomes in glioblastoma. Irish et al. developed an unsupervised and automated machine learning algorithm for risk assessment and population recognition to identify phenotypically distinct cell subpopulations (including antigen-presenting cells, various immune cells, endothelial cells, GB cells, etc.) and to determine whether these specific populations stratify the survival of GB patients (Luo et al., 2023).

Despite recent advancements in GB immunotherapy, the TME inhibits anticancer immune responses, leading to significant resistance. The integration of single cell sequencing with spatial transcriptome sequencing is crucial for examining the GB tumor microenvironment, hence promoting the advancement of effective immunotherapy (N. Zhang et al., 2021).

1.23 AI applications in drug discovery

AI can provide valuable findings into the intricate cellular and molecular crosstalk among cells that govern tumor behavior when applied to biomimetic models of the GB microenvironment. The integration of AI and 3D technology could have a profound influence on GB research. It can accelerate drug discovery, enhance personalized treatment approaches, and deepen our understanding of tumor biology (Bittner & Farajnia, 2022).

The pathophysiology of GB is complex, rendering it challenging to understand. Current targeted therapies often focus on experimental hypotheses while overlooking additional aspects, which can lead to potentially harmful effects (Luger et al., 2022). AI can address this complexity by comprehending gene interactions and product interactions, thus enhancing our understanding of GB etiology and investigating novel anti-GB targets (Kong et al., 2020).

The advancement of GB therapy faces obstacles in identifying novel molecules that can cross the blood-brain barrier (BBB), which protects the brain from blood-borne constituents and pathogens. High permeability is essential for efficient drug discovery (P. Chen, Wang, & Gao, 2024). AI predictive models can mitigate intricate, costly, and time-intensive tests. Although we have observed recent advancements, including robust ML methods for novel compound screening, graphite dot-assisted laser beam ionization mass spectrometry for comprehensive drug development, and high-content phenotype screening for pharmaceutical assessment, AI applications in GB medication discovery are still in their early stages (Issa et al., 2021).

Recent evolutions in AI have facilitated the creation of robust *in silico* techniques for elucidating intricate gene and gene-product interactions. Molecular interaction networks are progressively employed in machine learning to forecast physiologically significant interactions. SkipGNN, a graph neural network methodology, forecasts molecular cross-talks among cells by consolidating data from cell interactions, referred to as skip similarity. It gathers brain signals from both immediate and two-hop neighbors and non-linearly processes them to extract valuable information. Experiments on four interaction networks demonstrate enhanced performance, the acquisition of biologically relevant

embeddings, and proficiency in handling noisy, incomplete networks (Huang, Xiao, Glass, Zitnik, & Sun, 2020).

Graph neural networks (GNNs), including Graph Convolutional Networks (GCNs), Graph Attention Networks (GATs), and SkipGNN, have demonstrated remarkable efficacy in simulating biological networks by discerning patterns from protein–protein and gene–regulatory interactions (Yuan & Bar-Joseph, 2020). The research introduces EPGAT, a methodology for Essentiality Prediction utilizing GATs on graph-structured data. The model acquires gene essentiality patterns from protein-protein crosstalk networks and incorporates a multi-omics model represented as node attributes. EPGAT surpasses network-based and shallow machine learning methods, demonstrating competitive efficacy compared to the leading node2vec embedding technique. This is the most effective method in situations characterized by constrained and uneven training data (Schapke, Tavares, & Recamonde-Mendoza, 2022). Some research introduces a higher-order graph convolutional network (HOGCN) for predicting biological interactions. HOGCN consolidates data from the higher-order neighborhood, acquiring feature representations of neighbors at several distances. Experiments conducted on four interaction networks demonstrate that HOGCN yields more precise and calibrated predictions, particularly in the context of noisy, sparse networks. It also excels in literature-based case studies (Kc, Li, Cui, & Haake, 2022).

Non-coding RNA (ncRNA) plays a crucial role in biological interactions and mechanisms, which may be related to many human disorders. Identifying new ncRNA-protein interactions (NPIs) and mechanisms is crucial for comprehending the role of ncRNA. A novel computer model, named GATLGEMF, employs graph neural networks (GNNs) to forecast non-pharmaceutical interventions (NPIs). The model demonstrates exceptional performance on four benchmark datasets, attaining an AUC value of 92.41% for RPI2241 and 98.93% for NPInter v2.0 (J. Yan, Qu, Li, Wang, & Tan, 2024).

Another study presents a new EPI prediction approach, GATv2EPI, which utilizes a dynamic GNN. It utilizes epigenetic data on promoters and enhancers to investigate intricate regulatory patterns. The approach enhances EPI prediction precision by

encapsulating intricate topological structural information and emphasizing the significance of epigenetic characteristics (T. Zhang, Zhao, Sun, Gao, & Liu, 2024).

The Explainable Multilayer Graph Neural Network (EMGNN) represents an innovative methodology for cancer gene identification, utilizing various gene-gene crosstalk networks and comprehensive multi-omics data of cancer. This strategy surpasses current techniques, demonstrating an average enhancement of 7.15% in the precision-recall curve. EMGNN highlights newly predicted cancer genes that exhibit contradictory predictions across individual biological networks, providing significant biological insights through model-level feature importance determination and molecular-level gene set enrichment analysis. This new approach presents a robust new framework for cancer genomics research (Chatzianastasis, Vazirgiannis, & Zhang, 2023).

These models can elucidate both direct and indirect interactions among biological entities, enhancing prediction accuracy in sparse datasets and facilitating the identification of novel regulatory processes. Models such as EMGNN and GATv2EPI have demonstrated significant predictive capability in identifying cancer-related genes and enhancer-promoter interactions, with numerous predictions being confirmed and subsequently incorporated into experimental databases. BioGRID is an open-access database that houses curated protein and genetic interactions from several species, including yeast, worms, Drosophila, mice, and humans. Comprising approximately 1.93 million handpicked interactions, it facilitates biomedical discoveries, particularly in the context of human health and disease. BioGRID records protein post-translational changes and interactions with bioactive small compounds. It also engages in themed curation efforts focused on cellular regulation and disease domains. The Open Repository of CRISPR Screens (ORCS) aggregates single mutant phenotypes and genetic interactions obtained from genome wide CRISPR/Cas9 genetic screens (Oughtred et al., 2021). Simultaneously, Bayesian networks provide a probabilistic framework for integrating diverse biological data and managing uncertainty, particularly in the reconstruction of large-scale regulatory networks. Biological systems operate by intricate interactions among biomolecules, requiring a comprehensive, multi-omics approach. This necessitates the formulation of methodologies that encapsulate these connections and integrate diverse data. A significant challenge is the absence of

data, as not all biomolecules are quantified in every sample. Recent advancements in artificial intelligence and statistical learning have enabled the analysis of multi-omics data; nevertheless, many methods require access to completely observed data. This evaluation examines contemporary methodologies, conventional approaches to handling missing data, and prospective advancements (Flores et al., 2023). Collectively, these AI-based methodologies have yielded tangible results, including the identification of critical genes, prospective therapeutic targets, and previously unrecognized molecular interactions, thereby enhancing our understanding of disease causes and guiding precision medicine initiatives.

Our research incorporated machine learning models with image-based time-series data and experimental research centered on TMZ to forecast cellular behavior. The Materials and Methods chapter explains the experimental design, data collection protocols, and machine learning approaches utilized (CHAPTER 2, MATERIALS AND METHODS). The Results chapter delineates all calculations, model projections, and quantitative outcomes obtained from these methodologies (CHAPTER 3, RESULTS). The Discussion chapter interprets the data, examines methodological and biological limits, and situates the findings within the framework of previous literature (CHAPTER 4, DISCUSSION). The Conclusion chapter consolidates the main findings and outlines future perspectives and potential directions for further studies (CHAPTER 5, CONCLUSION).

1.24 Data management

Procuring adequate data for AI models remains challenging due to issues with data integrity and model training. Randomized controlled trials (RCTs) serve as the standard for training and verifying AI models; however, acquiring RCT datasets is labor-intensive and requires collaboration across multiple institutions to obtain extensive datasets (Wilkinson et al., 2016).

Federated Learning (FL) is an ML methodology that enables numerous devices or organizations to jointly train a model while preserving the confidentiality of raw data. Each participant trains the model locally using their own data and subsequently

transmits only the model changes, such as weights or gradients, to a central server, which consolidates these updates to enhance the global model. FL facilitates collaborative model training but necessitates secure aggregation to avert data leakage. A post-quantum aggregation technique utilizing a Key Homomorphic Pseudo-Random Function (KHPRF) is proposed, which minimizes communication cost and facilitates numerous aggregations with a single initialization, thereby achieving quantum-resistant aggregation with enhanced efficiency (Qin & Xu, 2025). The growing accessibility of healthcare data via computer software and hardware technologies offers a substantial opportunity for data science to enhance care delivery. Nonetheless, disjointed, and proprietary healthcare data hinders the ability to produce reliable outcomes across populations. Federated learning, a method that trains a collective global model via a central server while retaining sensitive data within local entities, presents potential solutions to these challenges and enhances privacy protection in the healthcare (Xu et al., 2021). Data-driven machine learning (ML) is a practical approach for constructing precise models from medical data. Access to this data is restricted owing to privacy issues and data silos (Rieke et al., 2020).

Organoid technologies have numerous applications in biomedical fields, including the creation of biobanks, developmental biology, precision medicine, drug screening, regenerative medicine, and disease modeling. They develop biomimetic 3D systems that mimic and operate like actual tissues, underscoring their significance for investigating. These applications may require particular data storage and management methodologies. Organoids maintain patient-specific biological variability, supporting extensive studies, and enable personalized therapeutic testing in precision medicine. Additionally, they enhance evaluations of drug efficacy and toxicity compared to conventional two-dimensional models, while also providing insights into tissue development and pathological mechanisms in regenerative medicine and disease modeling (Bai et al., 2024). The following figure (Figure 1.18) illustrates the applications of organoids in various domains.

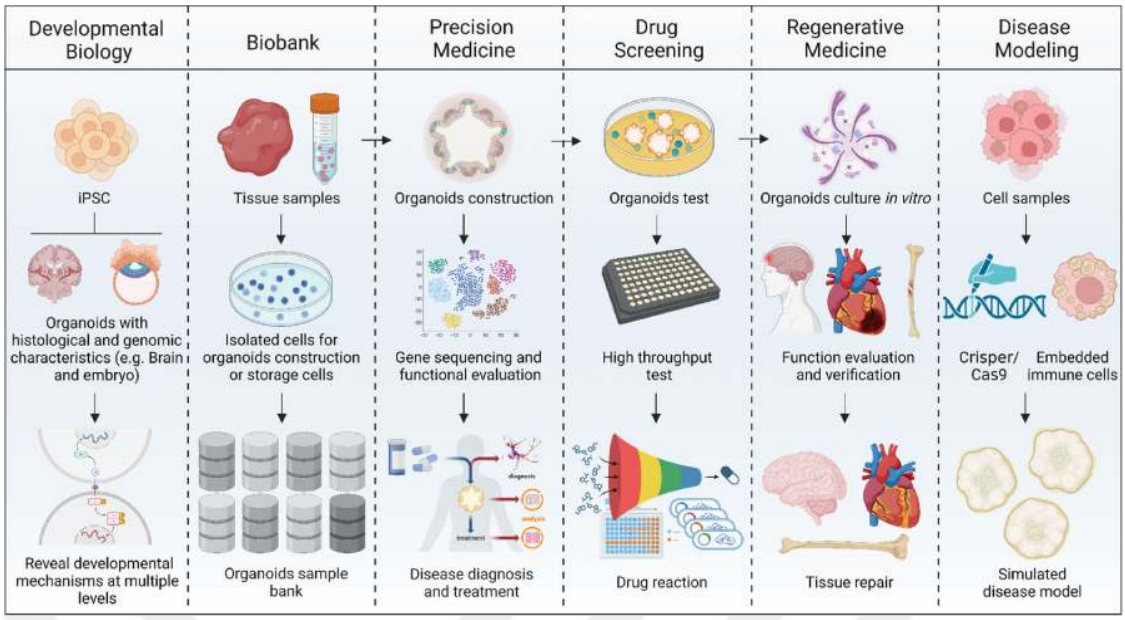

Figure 1.18: Applications of organoids (Bai et al., 2024)

Data gaps in AI-based organoid models are inevitable, as illustrated in the following graph (Figure 1.19). This challenge may be overcome by integrating multi-source data through novel methods, which could provide more insights into organoids (Bai et al., 2024).

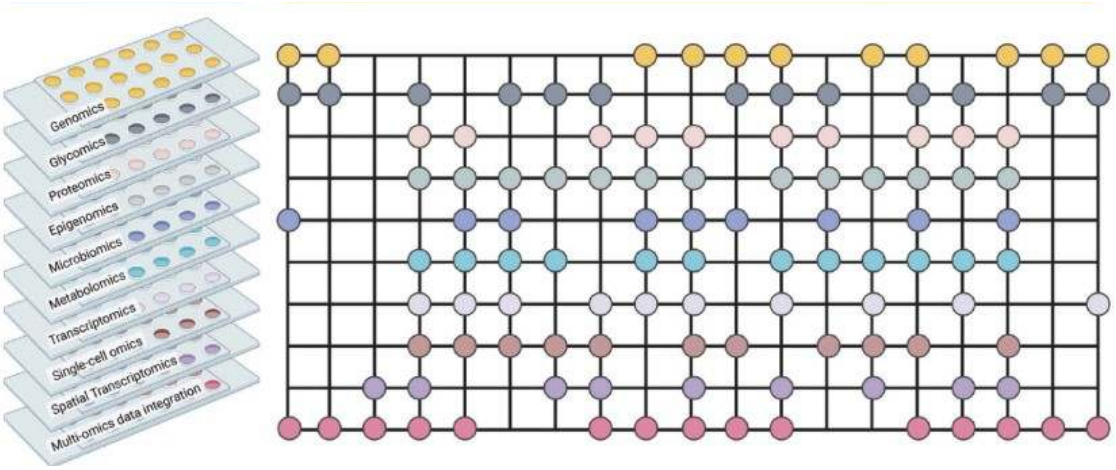

Figure 1.19: Data gaps in AI-based organoids in an omics perspective (Bai et al., 2024)

CHAPTER 2: MATERIALS AND METHODS

The methodology employed to address a problem significantly depends on one's perspective. Multidisciplinary studies play a vital role in finding solutions to complex problems. Mathematical and computer modeling is increasingly employed to clarify the complex dynamics of tumor formation in both animal and human settings. As the models develop, they are expected to predict patient outcomes and inform therapeutic protocols, including dosage and timing of administration. Furthermore, regarding biomimetic models, there is an opportunity to shift from animal models to more human brain-like constructs in *in vitro* 3D environments. he complexity of cancer, which includes genetic predisposition and the processes that allow abnormal cells to grow, invade, and migrate, has led to many problems that have not yet been solved. Employing a range of integrated methodologies may provide the most precise, predictive, reproducible, and efficient models to further neuro-oncological research (Praska, Odde, & Neil, 2021).

The Introduction chapter provides a thorough literature review to describe and contextualize the principal concepts discussed in this dissertation, along with extensive background information related to GB (CHAPTER 1, INTRODUCTION). GB cells have characteristic behaviors, including proliferation, invasion, and migration. In this research, software programming techniques are developed to calculate migration capacity and proliferation rate as parameter results. The study consisted of days three, five, and seven of two experiment sets; parameter findings were calculated using software programming approaches to analyze image files. The results dataset consists of two separate components, produced expressly for testing and training purposes, respectively.

All AI-based studies related to GB are often valuable, promising, and useful for patients and clinicians in facilitating their decision-making. Although AI applications range from disease severity diagnosis and prognosis prediction to treatment and postoperative

care, this thesis presents our study, which utilizes AI methods to predict tumor behavior using biomimetic 3D *in vitro* technologies as well as software design and programming methodologies.

Deep learning frameworks have significantly contributed to enhancing scientific studies in various fields, particularly biomedicine, by improving the accuracy of diagnoses, the efficiency of analysis, and the quality of decisions based on data. AI has evolved the way clinicians diagnose diseases, enabling faster and more accurate detection, the creation of personalized treatment plans, and enhanced patient care (Hinton & Jordan, 2024).

AI can process vast amounts of data and identify patterns more quickly than the human brain can. As a result, AI could be a valuable and powerful tool in neuroscience. In the field of malignancies, particularly regarding GB, AI has shown promise in enhancing diagnostics, predicting patient treatment responses, and identifying innovative therapeutic interventions. AI has become a valuable tool for scientific study because it facilitates the objective examination of large datasets, enables the identification of patterns and similarities, and facilitates the prediction of future outcomes (Al-Rahbi et al., 2024). Additionally, the application of AI algorithms in neuroscience research may enhance our understanding of brain functions and elucidate disease mechanisms (Macpherson et al., 2021).

Even with significant improvements in AI techniques, a clear gap remains between how computers are programmed and how the human mind makes decisions (J. Fan et al., 2020). AI refers to computer systems designed to mimic the mental processes associated with human reasoning. These systems operate using either rule-based logic or adaptive, data-driven algorithms, providing answers that are not limited by the rules established for them. Two core ideas are reviewed to have a deeper understanding of today's capabilities and potential applications of AI: Machine Learning (ML) and Deep Learning (DL) (Parekh & Jacobs, 2019).

The figure below (Figure 2.1) outlines a summary of the key ideas from the perspective of AI applications in this dissertation. In literature, developing a DNN for cell counting and classification is feasible (Jiang et al., 2021). Similar to the method explained in the

following figure (Figure 2.1), our research project includes image datasets collected on various days, indicating the presence of time-series data. ML algorithms are well-suited for predicting the future activity of cells, whereas CNNs are ideal for extracting spatial features from images. RNNs (especially LSTM/GRU) handle temporal sequences. The combination (CNN–RNN hybrid) is most accurate for spatiotemporal data such as time-lapse microscopy. CNNs process spatial image features while RNNs model temporal evolution. Sequence learning models, including RNN, LSTM, and GRU circuits, are illustrated in the following figure (Zargar, 2021).

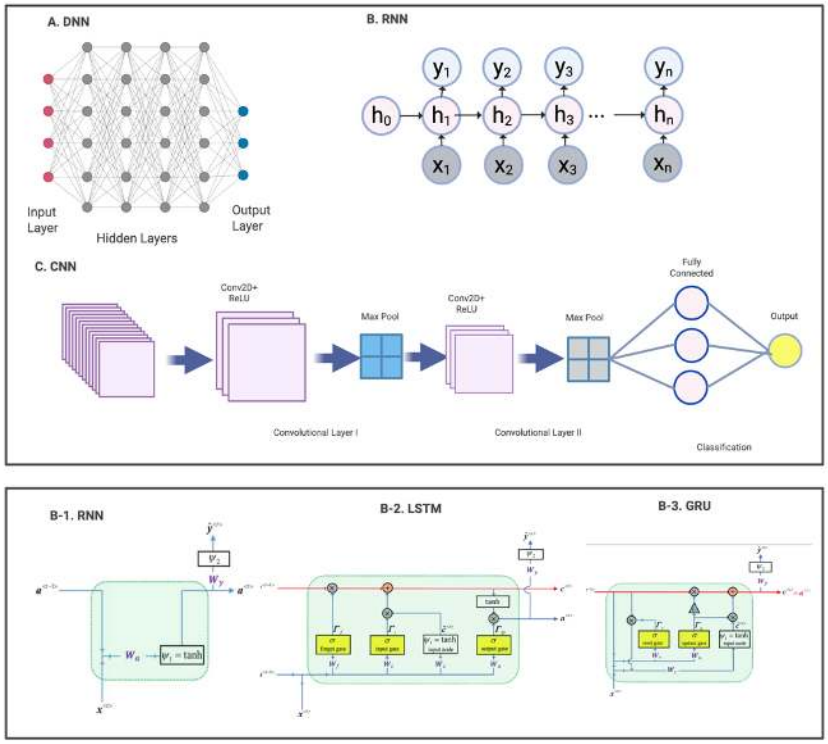

Figure 2.1: Brief outline of the fundamental concepts in AI. Definitions arise for two fundamental terms: machine learning and deep learning. CNN denotes convolutional neural networks, RNN signifies recurrent neural networks, and DNN represents deep neural networks (Luo et al., 2023). Sequence learning models, including RNN, LSTM, and GRU circuits, are illustrated in B-1, B-2, and B-3 (Zargar, 2021).

The main objective of this study is related to applied physics, as it simulates the physical properties of cells in 3D multiplex systems. The method is to utilize computational tools and ML algorithms on 3D cell culture images, as today we observe a lack of time-series predictive models on 3D GB datasets. By filling this gap, we aim

to gain a deeper understanding of TME and discover optimal solutions for its treatment. Combining these advanced methods may be beneficial for scientists to make accurate predictions about how cells will evolve and behave over time. These improvements may also facilitate the discovery of new parameters and therapeutic targets, leading to more precise medicine for patients over time. Additionally, our results may pave the way for subsequent investigations that delve deeper into the complex nature of tumor dynamics within 3D multiplex environments, allowing for the simultaneous capture of a large amount of data.

The methodology utilized in this research encompassed the following elements:

1) 3D cell culture in multiplex tumor microenvironments.
2) Getting time-series image datasets from confocal microscopy.
3) Using traditional computational methods to calculate parameters from image datasets.
4) This research used the transformers as an ML model to predict GB cell behaviors.

On the other hand, in addition to the core of this investigation, we enriched our research by measuring the proliferation behavior of live cell imaging videos captured for 3 days. Another application was the comparison of experimental and control studies of the control (CTRL) and TMZ groups. For this part of the study, we used a linear regression model to predict upcoming results.

5) Using ML algorithms to predict cell behavior for the comparison of CTRL/TMZ results.
6) Observing cell behavior on live cell imaging and voxel-based volume calculation.

2.1 3D cell culture on TME multiplexed platforms

In this study, RFP-tagged human brain U87 GB cells and GFP-tagged I-NHA human brain astrocytes were cultured in GelMA-based hydrogels on 3D multiplex platforms, which enable the culturing of biological replicates. Cell metabolic activity at distinctive

concentrations (1 million cells/mL, 2 million cells/mL, 3 million cells/mL) was quantified using the MTT assay (Ismayilzada, 2024). Figure 2.2 demonstrates that U87 cells exhibited a similar pattern, with a concentration of 2 million/mL indicating a fourfold increase in viability that was significantly higher than both 1 million/mL and 3 million/mL concentrations ($p = 0.0139$ and $p < 0.0001$, respectively). We were able to recreate a cell culture environment in this gel to study how GB and astrocytes interact and how this interaction changes over time. We obtained over 300 cell images. When we consider days 3, 5, and 7 as a combined set of image data, we have approximately 70 sets of image data, which may be beneficial as input data for ML models predicting cell behavior.

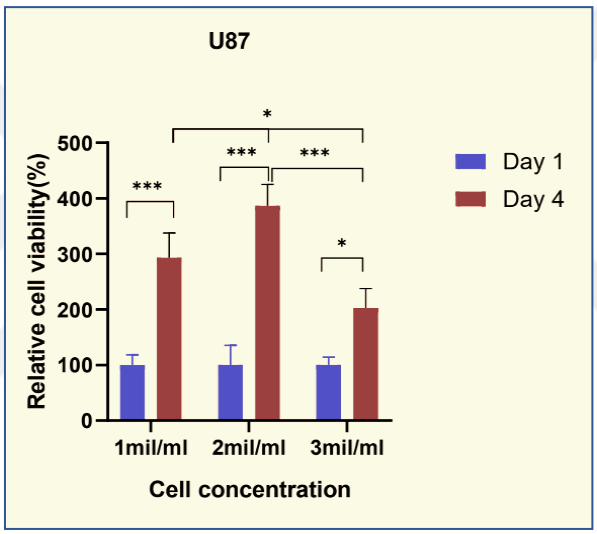

Figure 2.2: Human brain U87 GB cell metabolic activity at distinctive concentrations (1 million/ml, 2 million/ml, 3 million/ml), quantified with MTT assay

Advancements in biomimetic 3D cell culture platforms, such as GelMA-based hydrogels and bio-printed structures, have facilitated the recreation of physiologically relevant GB microenvironments, enabling enhanced evaluation of invasion and therapeutic response (Ismayilzada, 2024). GelMA was produced through a one-pot method and assessed via rheological tests and scanning electron microscopy to confirm its tunable stiffness and porosity. Using a two-step photolithography method, micropatterned co-culture constructs were created that allowed human brain U87 GB cells and I-NHA human brain astrocyte cells to be placed precisely where they were needed, either in direct contact with GelMA microspheres or spread out over nearby astrocytic matrices. We used MTT and Live/Dead tests to verify whether the cells were

alive, and phalloidin staining to observe the organization of the cytoskeleton. Functional analyses included RNA extraction, cDNA synthesis, and RT-qPCR targeting genes related to invasion, hypoxia, and epithelial-mesenchymal transition, as well as immunofluorescent labeling of GFAP to evaluate astrocyte reactivity. In addition, human brain U87 GB cells and I-NHA human brain astrocyte cells were labeled with RFP/GFP tags for imaging. MATLAB was used to obtain confocal microscopy images to determine how well they could move based on the normalized pixel area outside the microsphere limits, while calculating parameters including migration capacity and proliferation rate. Our research demonstrated that a biomimetic 3D multiplexed cell culture environment facilitated the simultaneous recording of many data points under identical conditions. The following metrics can be derived from datasets; this research emphasizes migration capacity and proliferation rate as calculated parameters. These parameters can be defined separately by researchers conducting the study. Below are the definitions in the context of this investigation; the ones that we have focused on are migration capacity and proliferation rate metrics.

- **Proliferation Rate**: Total area of U87 cells, which are RFP tagged red pixel points in the image dataset. It is also possible to calculate this metric as a volume if the image has all its dimensions or slices.
- **Cell Morphology**: U87 cells have a spherical shape, characterized by their radius; other types of GB cells, such as those other than U87, may have separate structures, including a diamond shape.
- **Migration Capacity**: The percentage of the total number of U87 cells that have moved outside the donut compared to the total area of the frame (1024x1024).
- **Invasion Rate**: This parameter can be defined separately according to perspective; in our study we define it as the number of surrounding micro glial cell types which can be astrocytes or Brain Microvascular Endothelial Cells (BMEC) cells that are co cultured with U87 GB cells and the total area of micro glial cells that have invaded inside to the donut represent invasion rate, it is measured by total pixel area of micro glial cells, which were GFP tagged and had green color on image datasets.
- **Migration Rate**: This parameter represents the total number of U87 cells that migrated outside the designated area, specifically outside the donut. It is

measured by the total pixel area of RFP-tagged U87 cells, which appear outside the green donut in the image dataset.

We had two particular types of co-culture set-ups: direct and stamping co-culture configurations. In direct co-culture, GB cells and astrocytes are grown together in a donut-shaped environment, as shown in the following figure (Figure 2.3). In the stamping configuration, on the other hand, GB cells are placed inside the donut and surrounded by astrocytes. We observed that the migration capacity in the direct setup significantly surpasses that of the stamping arrangement. Therefore, we decided to use a direct arrangement for the part of our study that involved making predictions during the computational part of this investigation.

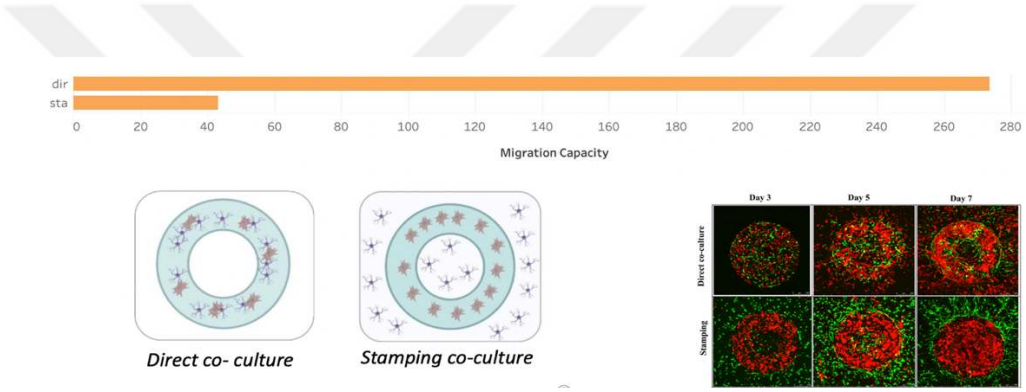

Figure 2.3: Migration behavior of U87 GB cells, comparison of direct and stamping configurations for migration capacity

2.2 Data acquisition and preprocessing

Confocal image datasets were acquired with a spectral confocal laser scanning microscope, Leica Microsystems (Ismayilzada, 2024). The software can provide images in both 2D and 3D formats, as well as create videos for visualizing the image from distinct perspectives. Our study included confocal images taken on the following days: 3, 5, 7, and 9. A set of samples is illustrated in the following figure (Figure 2.4), which presents both direct and stamping configurations in a time series format. Due to the absence of data for day 9 in the second biological duplicate, we incorporated only the photos from days 3, 5, and 7 in our calculations and predictions within the computational segment of our research. The results will be expanded and evaluated in the subsequent chapters (RESULTS, DISCUSSION).

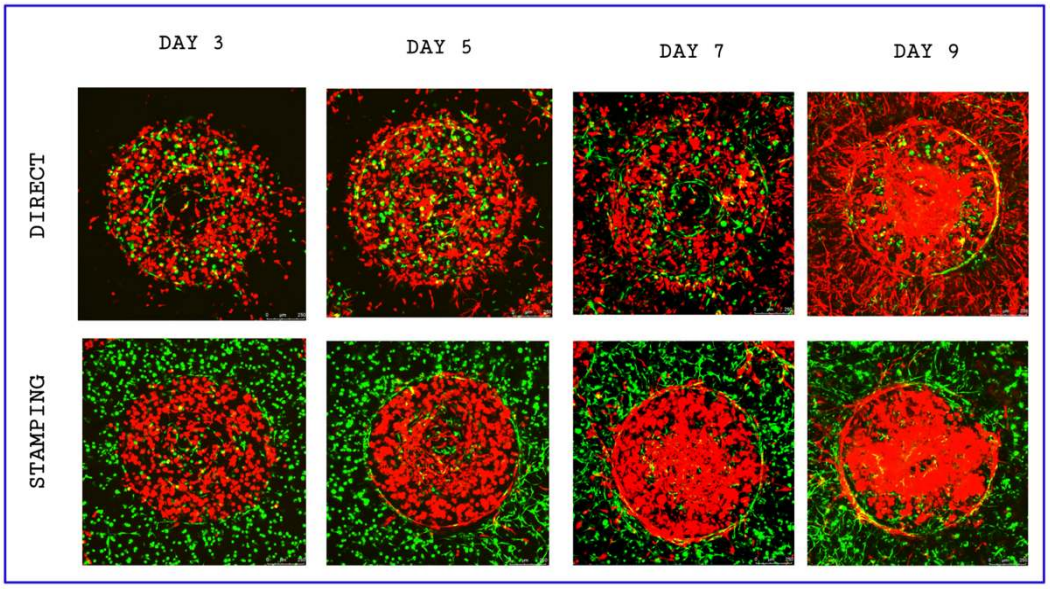

Figure 2.4: 3D confocal microscopy images of RFP-tagged human brain U87 GB cells and GFP-tagged I-NHA human brain astrocytes from day 3 to day 9 in direct and stamping configurations (Ismayilzada, 2024)

The fluorescent microscopy images in the following figure (Figure 2.5) demonstrate RFP-tagged human brain U87 GB cells on the 3rd, 4th, and 14th days of cultivation. On the 3rd day, RFP-tagged U87 cells cultured in a 2D environment with GB medium developed a neurosphere-like structure. On the 4th day, monoculture U87 cells are observed and demonstrated at distinct zoom levels. On days 3 and 14, cells are treated separately to form spheroids. Ultimately, we observe a clear proliferation and migration of RFP-tagged U87 spheroid cells cultured in a 2D environment with GB medium at day 14.

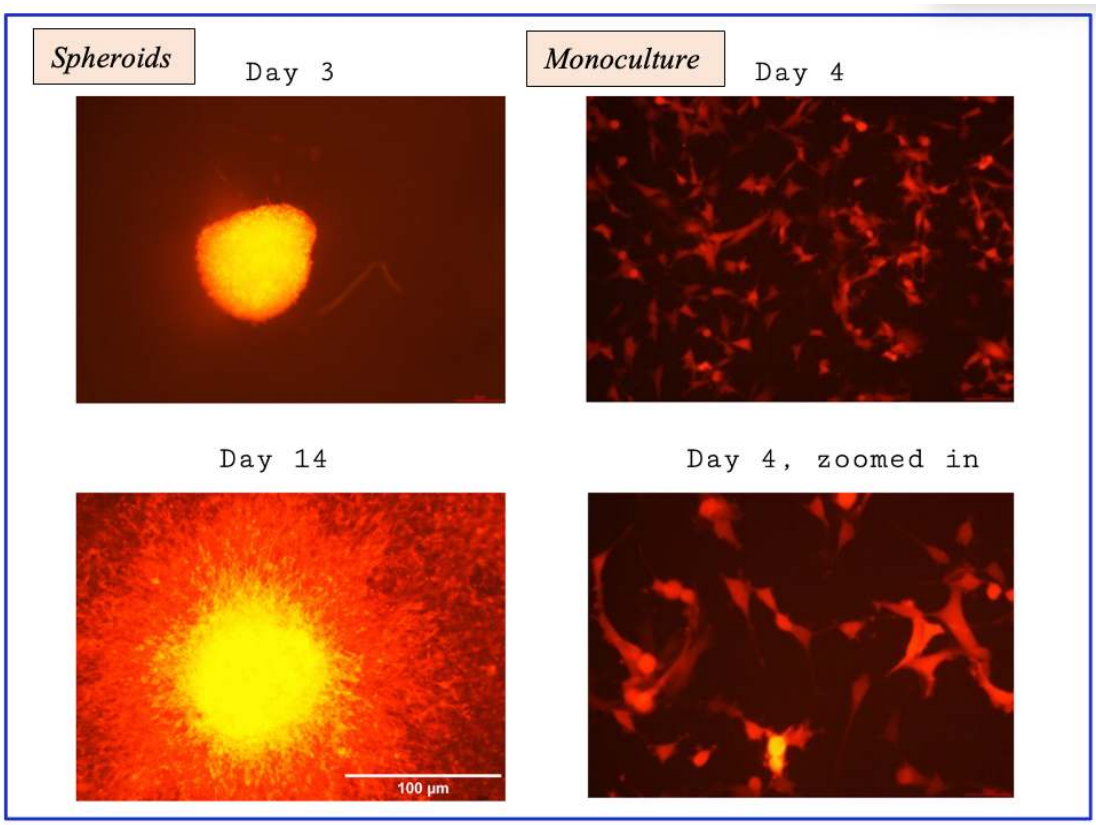

Figure 2.5: RFP-tagged human brain U87 2D fluorescent microscope images

In the following figure (Figure 2.6), inverted microscope images of the human brain U87 are shown, comparing both the control and TMZ treatment groups (Oğuz, 2025). Images were taken at three distinct times: at the beginning, after 24 hours, and after 48 hours, for each case. This time-lapse display enables a qualitative evaluation of cellular behaviors, encompassing proliferation, morphological changes, and migration dynamics. The comparative imaging among treatment groups elucidates the temporal effects of TMZ on human brain U87 GB cell proliferation and migration.

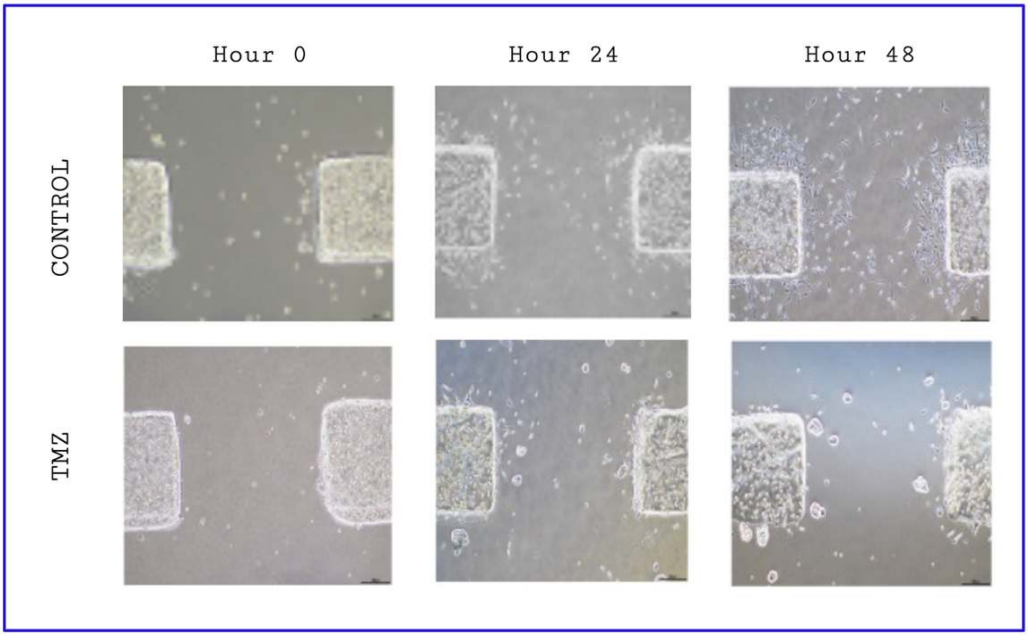

Figure 2.6: Time-based TMZ-Control U87 Inverted Microscopic Images (Oğuz, 2025)

2.3 Live cell imaging

We can observe how cells evolve both in 2D and 3D platforms. A case study examines T-cell behavior via live cell imaging techniques. BEHAV3D is a technology that advances immuno-oncology research by integrating multi-color live 3D cell imaging and computational analysis, facilitating real-time observation of patient-derived tumor organoids and immune cells. It elucidates essential processes, including tumor cell apoptosis and T-cell migration, using rapid imaging methodologies, thereby facilitating comprehensive studies at both the single-cell and population scales. The platform's image segmentation algorithms generate data-driven 3D renderings that highlight tumor heterogeneity, demonstrating its efficacy as a powerful tool for evaluating immunotherapies in diverse co-culture settings (Alieva et al., 2024).

Cell migration is essential for accurate cell localization and function during development and homeostasis, with heightened migratory potential noted in disorders such as cancer. Another study presents a technique for prolonged 3D imaging using two-photon excitation microscopy, which facilitates an understanding of cellular dynamics in both healthy and diseased tissues. Refined live imaging techniques were

established to examine cell migration (Staneva, Barbazan, Simon, Vignjevic, & Krndija, 2018). The following figure (Figure 2.7) illustrates the application of the Otsu thresholding method to the first four frames of the live cell imaging sequence, enhancing contrast and facilitating quantitative segmentation. In our research, live cell images were acquired with Zeiss confocal microscope.

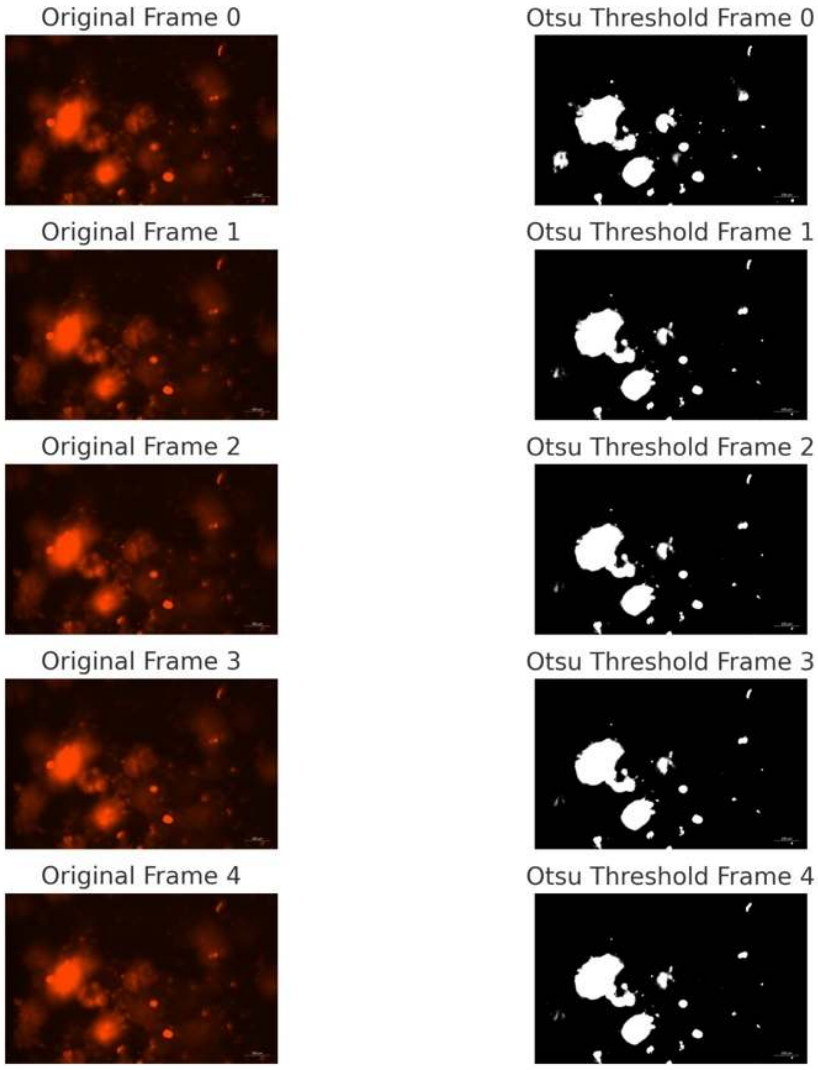

Figure 2.7: Otsu thresholding on live cell imaging frames

2.4 Processing data

In this research, we utilized MATLAB and specific Python (version 3.1) development libraries, including TensorFlow, to assess the migration capacity and proliferation rate of U87 tumor cells. This was done by examining images taken with a confocal

microscope. The pictures were taken on the following days: 3, 5, and 7, and the proliferation rate gave an idea of the growth of tumor cells. Before machine learning algorithms were used, traditional software programming methods were used to determine and calculate parameter outcomes from image datasets (Prakash & Kanagachidambaresan, 2021). The results dataset had two distinct components, explicitly created for testing and training purposes in the ML process. The TensorFlow software package was utilized to develop RNN, LSTM, and GRU neural networks (Pavithra, Saruladha, & Sathyabama, 2019), which are kinds of ML algorithms.

In MATLAB, paired visualizations can be developed as in the following figure (Figure 2.8), allowing for direct image comparison. This feature enables researchers to compare original and processed images or particular types of cells side by side, facilitating a more accurate evaluation of cell interactions. Such comparative visualization is highly beneficial for cell culture image analysis, where precise analysis of variations in segmentation is required.

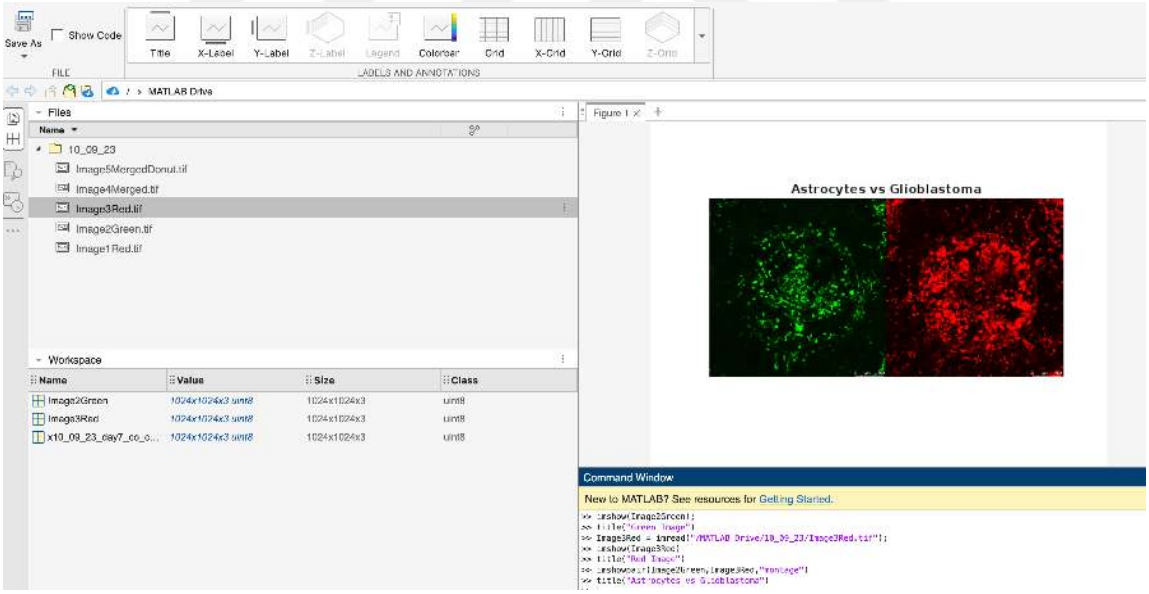

Figure 2.8: Paired comparison of confocal images on MATLAB, GFP-tagged I-NHA human brain astrocytes vs. RFP-tagged human brain U87 GB cells

Thresholding is an essential phase in image segmentation, with studies demonstrating a new approach to utilizing 3D Otsu and multi-scale representation for segmenting medical images. A faster version of this method is presented to reduce its high temporal complexity by dimension decomposition. The method employs a multi-scale

representation to reduce noise and enhance weak edges. It works by splitting images into smaller pieces using 3D Otsu and then smoothing them using quick local Laplacian filtering before the next iteration. Majority voting is used to get the final segmentation results. The algorithm is stable, capable of managing noise, and proficient in both dual and multi-level thresholding. Evaluations of MR brain pictures indicate that it surpasses alternative multi-level thresholding techniques (Feng, Zhao, Li, Zhang, & Li, 2017).

When we applied Otsu thresholding in MATLAB to our dataset, the segmentations resembled those in the following capture (Figure 2.9). MATLAB offers an online, accessible, and user-friendly interface for adjusting parameters, allowing researchers to instantly modify thresholding settings and observe their impact on image quality. This flexibility is especially beneficial when interacting with cell culture images that are not all the same, as the optimal threshold values may vary for each sample and require refinement over time. However, in our study, we applied the same threshold value determined by Otsu's method throughout the dataset to ensure consistency across all samples.

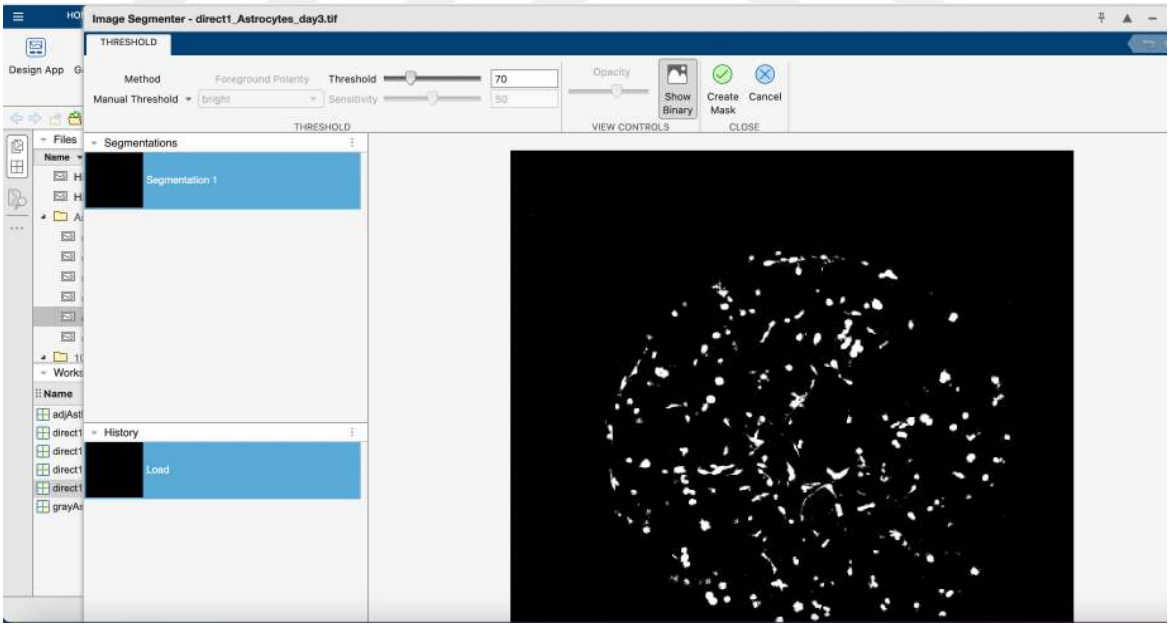

Figure 2.9: Otsu thresholding on MATLAB interface

In addition to the MATLAB programming language, we also used Python (version 3.1) for Otsu thresholding. The following images (Figure 2.10) were taken to observe the growth of U87 GB tumor cells; the frame numbers represent the 1st, 16th, 32nd, and

48th hours after the samples were placed in the microscope. Cell growth can be calculated after applying the Otsu thresholding algorithm.

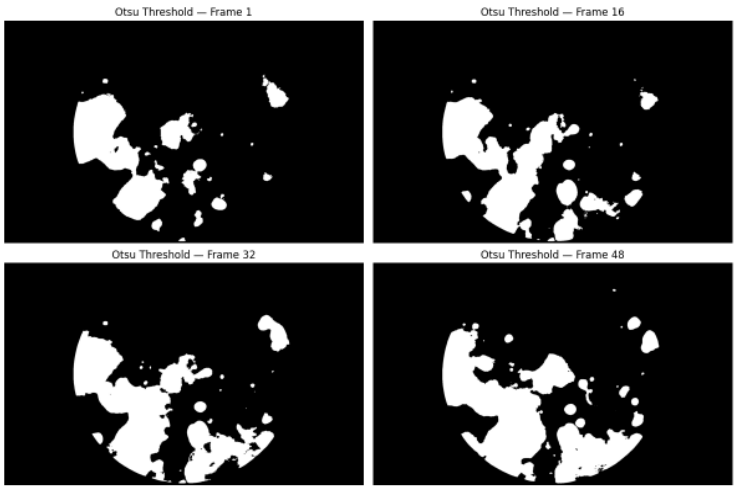

Figure 2.10: Otsu thresholding on Python, hourly GB cell growth

This document summarizes the parameters and their calculations.

- *Migration Capacity*: The proportion of U87 cells that moved from the interior to the exterior of the donut relative to the overall exterior area of the frame (1024x1024, pixels).

$$\text{Migration Capacity} = \text{Area in Pixel Net} * 100 / \text{Area Exterior the circle} \tag{2.1}$$

- *Proliferation Rate*: The total area occupied by U87 cells, represented as red pixel points labeled with RFP in the picture collection.

$$\text{Proliferation rate} = \text{Area in pixels} \tag{2.2}$$

The future behavior of cells and the calculated metrics were predicted using ML techniques, with Python utilized as the programming language in an Anaconda environment (Müller & Guido, 2016). Migration capacity, as a metric of GB, indicates the percentage of the total migrated U87 cells relative to the entire area of the frame (1024 x 1024). Data were expressed as the mean ± standard error (N=2).

2.5 Volume calculation on 3D cell culture images

The following figure (Figure 2.11) illustrates a 3D confocal microscopy configuration of RFP-tagged U87 human GB cells that are surrounded by GFP-tagged human BMEC cells. This 3D image illustrates the spatial arrangement of the two groups of cells on day 7 in the stamping configuration. It additionally demonstrates how closely they interacted with each other and how they were structured within the co-culture system. High-resolution confocal imaging enables a detailed examination of how tumor and endothelial cells interact, providing valuable insights into how tumors invade and how the microenvironment responds in tailored *in vitro* models.

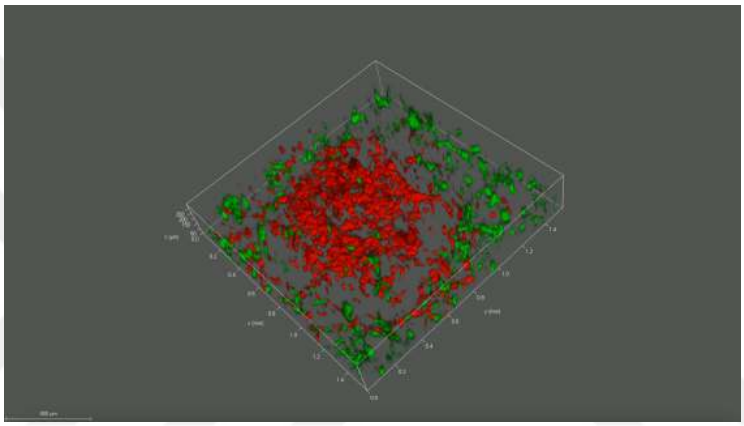

Figure 2.11: 3D confocal image of U87 and BMEC cells in stamping configuration, day 7, scale 500 µm

The following image (Figure 2.12) represents a 3D confocal microscopy output for RFP-tagged U87 GB cells surrounded by GFP-tagged BMEC cells at day 14.

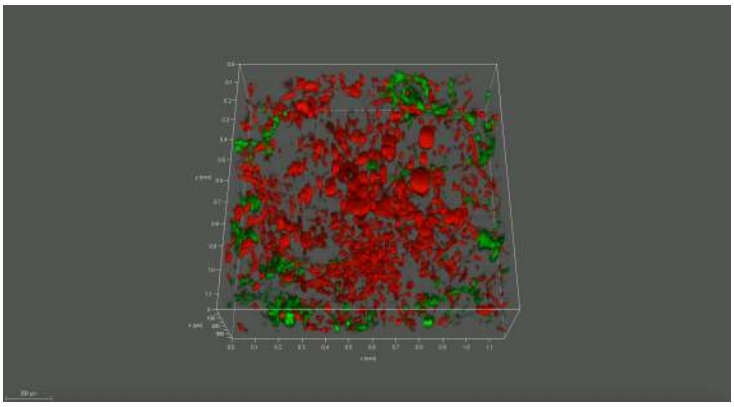

Figure 2.12: 3D confocal image of U87 and BMEC cells in stamping configuration, day 14, scale 200 µm

3D confocal images encompass three spatial dimensions and defined scaling, which enables the estimation of the volume of formations, such as tumor cell aggregation. The table below includes the pseudocode that explains how the volume of RFP-tagged U87 cells was calculated.

First, the 3D image stack was imported, and the X, Y, and Z axes ere defined for volumetric analysis. After converting the images to RGB format, HSV thresholding with a specified hue range was applied to identify regions of red fluorescence corresponding to tumor cells. A binary mask was then generated to isolate red pixels. These pixels were summed across the X - Y plane for each slice to calculate the tumor area per slice.

For normalization, the total number of pixels (voxels) in the image was used, allowing the percentage of the area occupied by tumor cells in the top and bottom slices to be determined.

An approximate height was then estimated from the extremity slices along the Z axis. Finally, tumor volume was calculated by multiplying the average cross-sectional area by the estimated height. This voxel-based approach provides a consistent method for quantifying tumor volume in 3D confocal microscopy datasets.

The following table (Table 2.1) summarizes the steps of this algorithm in pseudocode format.

Table 2.1: Pseudocode of the volume calculation algorithm utilizing 3D confocal microscopy images based on voxels

```
Input: 3D Confocal Image
Output: Volume in terms of micrometer square of the U87 Brain
tumor cells

1. Load the image
2. Define X,Y,Z axis for calculating the volume
3. Convert to RGB to BGR
4. Extract red regions using HSV thresholding
5. Define red color range (two ranges to capture full hue wrap-
around)
6. Define a mask
   if the pixel is red
         mask_value <- 1
7. Calculate area in pixels
   For each pixel in X axis
      For each pixel in Y axis
            if mask value > 0
                  red_pixels <- sum(mask_value)
            endif
      endfor
   endfor

8. Calculate total pixels for normalization
   total_pixels <- area of the frame of X-Y axes
9. Calculate percentage area via top and bottom slices
   red_area_percent = (red_pixels / total_pixels)

10. Calculate approximate height via Z axis via 2 extremity
slices
11. Calculate overall best estimated tumor voxel volume
    Voxel_volume <- Average Area * average height
```

Colab functions as a platform for executing the Python programming language and offers the primary environment for software implementations. Colab, a cloud-based programming interface operating directly in the browser, utilizes Google Cloud resources to provide users with instant access to a fully configured Python workspace, complete with allocated processing and memory capacities. This configuration enables practical experimentation, programming, and cooperation, rendering it a valuable platform for all those involved *in silico* studies. It enables the deployment of code on GitHub, a platform that facilitates the creation, storage, management, and sharing of scripts by developers. The following capture (Figure 2.13) demonstrates the

implementation of voxel-based volume calculation on the Google Colab interface with Python (version 3.10).

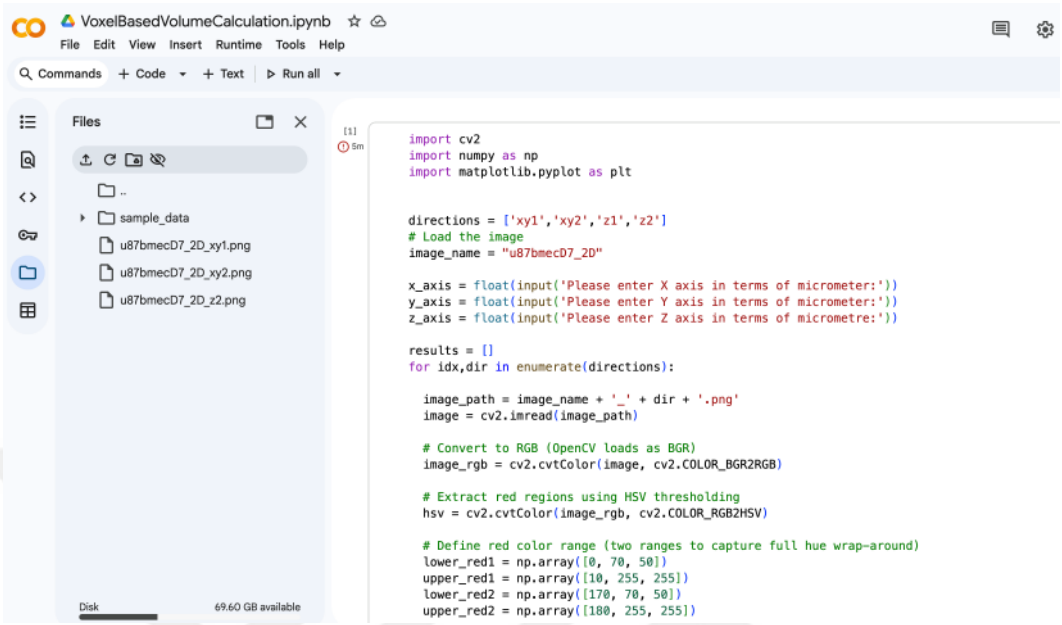

Figure 2.13: Voxel-based volume calculation on the Google Colab interface with Python programming language

2.6 Software programming for parameter calculations and predictions on time series data

The computation of the subsequent components is essential for determining the metrics associated with GB growth; we refer to them as parameters or metrics in this research. Metrics for the photos were calculated using the structured programming language MATLAB, employing area and radius calculations (Fangohr, 2004).

- Area (μm²): Gross/Net Area
- Diameter/Radius (μmeter): Estimated diameter of the spheroid

The cell images captured on days 3, 5, 7, and 9 are presented in the subsequent image dataset. Red U87 type GB cells engaged with green astrocytes, as in the following figure (Figure 2.14) (Ismayilzada, 2024). In the image below, the radius of the donut shape is visible on day 3. The images include both GFP-tagged I-NHA human brain astrocytes and RFP-tagged human brain U87 GB cells. It is visible to the human eye

that GB cells tend to grow from day 3 to day 9, and tumor cells migrate outside the donut.

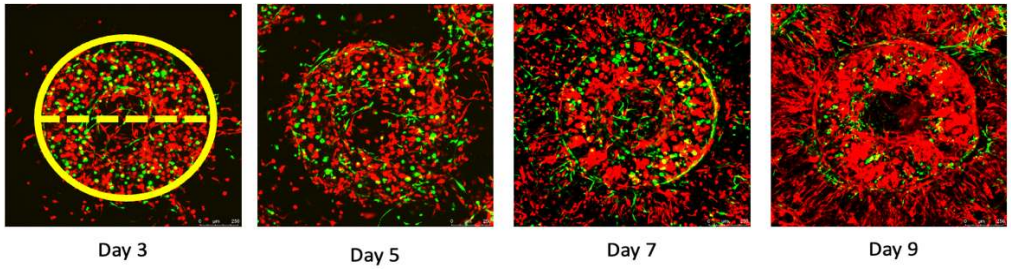

Figure 2.14: Calculating parameters based on confocal images (Ismayilzada, 2024)

The future behavior of cells and the computed metrics were forecasted using machine learning methods, with Python employed as the programming language within an Anaconda environment (Müller & Guido, 2016). Executed MATLAB Results for migration capacity parameters and a comparison of direct and stamping co-culture; these configurations will be incorporated as features in ML algorithms for future predictions of cellular activity. Migration capacity, as a metric of GB, indicates the percentage of the total migrated U87 cells relative to the entire area of the frame (1024 x 1024). Data were expressed as the mean ± standard error (N=2).

In a general perspective, time series data refers to a succession of observations recorded over time, with each data point associated with a specific timestamp (Kostadinov, 2018). We examined confocal images taken on days 3, 5, and 7.

The normalization process used MinMax functions before ML algorithms were used. RNNs provide a distinctive advantage in time series forecasting by facilitating the understanding of complex temporal relationships. Examining LSTM and Gated Recurrent Unit (GRU) models demonstrates to us how powerful predictive analytics can be considered as an ML approach (Zaccone & Karim, 2018).

In our study, we utilized several software platforms for programming purposes, including Google Colab, PyCharm, and Visual Studio Code. As shown in the following capture (Figure 2.15), PyCharm provides a user-friendly interface for performing tasks

by utilizing the client's (i.e., personal computer) local memory and processing resources.

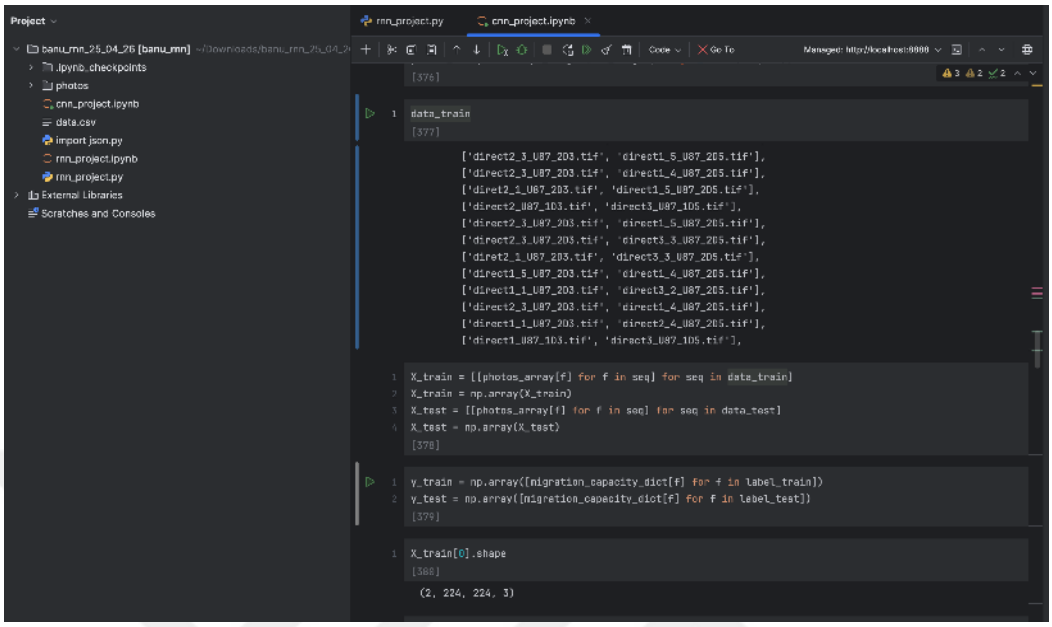

Figure 2.15: Neural networks project development on the PyCharm interface with the Python programming language

The loss function plays a crucial role in training various neural network architectures, including CNNs, RNNs, LSTMs, and GRUs. The loss function measures the accuracy with which the model's predictions align with the target outputs. This information is then used directly in the optimization techniques through backpropagation approaches, which offer a quantitative assessment of prediction inaccuracies, facilitating weight adjustments during training. The distinctions in architecture determine how gradients spread and how they affect the learning process. The loss, on the other hand, remains the primary goal of optimization, which aligns model predictions with the desired outputs. In CNNs, the loss function measures the discrepancy between the predicted labels and the actual labels. For example, cross-entropy loss is commonly used for classification, while mean squared error is typically employed for regression. In RNNs, the loss function examines predictions at each time step and compares them to the ground truth. This captures temporal relationships. LSTMs alleviate the vanishing gradient problem by allowing gradients to propagate over more extended periods of time. GRUs employ a simpler gating structure yet still exhibit disparities between sequences and adjust their weights.

We used the mean absolute percentage error (MAPE) to validate the accuracy of the prediction models. A lower percentage error generally means that the model can make more accurate predictions. Two evaluation settings were used to assess the model's accuracy. Training accuracy was measured using the same image data for both the training and testing phases, while test accuracy was determined by using the trained model on separate images not included in the training phase. We built three machine learning models to compare them. The first model combined a CNN with a long short-term memory (LSTM network; the second model employed RNN architectures, including LSTM and GRU layers, which are subtypes of RNNs. The third model combined CNN and LSTM elements again, which is why it can be referred to as a hybrid algorithm. We used the mean absolute error of percentage (MAEP) to measure the performance of the CNN-LSTM model, as it provides continuous regression values. Thus, overall accuracy did not apply to this architecture.

Accuracy is calculated as in the following equation, where N denotes the total number of samples with both predicted values and ground truth. In this formula, $\hat{y}_i$ is the model's predicted label for the i^{th} sample, and y_i is the true label for that sample. The indicator function $I(\hat{y}_i = y_i)$ is equal to 1 when the predicted label is the same as the true label and 0 when it isn't. Thus, the logic only counts correct predictions.

$$Accuracy = (1 / N) \cdot \Sigma\, I(\hat{y}\_i = y\_i) \tag{2.3}$$

The following equation, MAEP, is used to assess the prediction performance of ML models, including the CNN-LSTM model, $|\hat{y}_i - y_i|$ signifies the absolute error.

$$MAEP = (1 / N) \cdot \Sigma\, |y\_i - \hat{y}\_i| \tag{2.4}$$

In our study, we applied RNN, LSTM, and GRU as ML models and compared how accurate they were. We found that all three models functioned accurately; however, the GRU model yielded the most accurate results in terms of both accuracy and efficiency on the dataset we used. This suggests that GRU may be the best choice for our dataset and experiment design when it comes to tasks that require predicting time series sequences from cell culture images.

2.6.1 RNN

An RNN is a specialized kind of ML algorithm that can handle time series or sequence data. RNN can be considered as a smart cell with its own memories. The RNN understands each word while keeping track of what came before it in its memory. This memory enhances the algorithm's ability to identify patterns and connections in sequences, enabling the prediction of the next word. RNNs are practical algorithms for predicting future values in time series data, including stock prices or weather conditions, when historical data is available (J. Wang et al., 2022). In our study, we applied an RNN over an image dataset taken on distinct days; therefore, we consider the dataset as time series data.

RNNs are ideal ML models in terms of making predictions in time series data, as they offer a lot of benefits. They can handle sequential data of distinct lengths and perform accurately to capture long-term dependencies and trends over time. RNNs can handle time intervals that are not evenly spaced and adapt to distinct forecasting requirements by utilizing input and output sequences with varying time slots (Gulli, Kapoor, & Pal, 2019).

However, RNNs have issues such as the disappearing or vanishing gradient problem, which makes it harder for the algorithm to find long-term dependencies, considering that RNNs can be unraveled in this context. Memory constraints might also hinder their performance with excessively long scene sequences. LSTMs and GRUs are effective in solving certain types of problems, but in more complex time series situations, more advanced architectures, such as transformers, may work more accurately than RNNs. This means researchers or programmers must be cautious when choosing a model, as its accuracy may vary depending on the ML algorithm (Gulli et al., 2019).

The vanishing gradient problem makes algorithms hard to train DL networks, especially RNNs. This occurs when gradients, which indicate the magnitude and direction of weight changes during training, become increasingly smaller as they propagate backward through the layers. Such an issue complicates RNNs' ability to acquire long-

range dependencies, resulting in prolonged or less effective training. These ML methods not only improve the training process but also enhance the overall performance and accuracy of neural network models. By applying these solutions, practitioners can efficiently use DL architectures, allowing them to address more complex challenges with improved efficiency. The vanishing gradient problem is especially challenging for sequences that need long-term information storage or sharing, as it makes it harder for the network to find pattern similarities. Several ideas have been proposed to address the vanishing gradient problem, which reduces the effectiveness of training neural networks. To maintain stable gradient flow and prevent early failure, it is advisable to employ techniques such as proper weight initialization, batch normalization, gradient clipping, skip connections, and learning rate scheduling. These strategies facilitate training and enhance the overall accuracy of deep learning models. When practitioners employ these strategies, they can enable their networks to perform with high accuracy in handling complex tasks that involve sequential data. Among these several methodologies of RNN, two prominent solutions emerge, LSTM and GRU, which are subtypes of the RNN (Tolstoluzka & Telezhenko, 2024).

2.6.2 LSTM

The concept of Long Short-Term Memory was released in 1997, and this ML model effectively addresses complex, synthetic, long-time-lag tasks that previous recurrent network methods have failed to find solutions (Hochreiter & Schmidhuber, 1997).

The vanishing gradient problem is challenging for classical RNNs; finding long-term dependencies in sequential data is not effective enough in these ML models. However, LSTM effectively solves this problem by utilizing special memory cells and restricting mechanisms to maintain a constant flow of gradients over long sequences of data. This contributes to the network's understanding of long-term dependencies and significantly enhances its ability to learn from data that arrives sequentially. LSTM has three gates: the input gate, the forget gate, and the output gate. It is a kind of algorithm, and remarkably effective at capturing long-term dependencies (Tolstoluzka & Telezhenko, 2024).

2.6.3 GRU

LSTM and GRU are improved versions of RNNs that address the challenges inherent in RNNs, including the vanishing gradient problem. It can be used to detect the progression of certain diseases. GRU is a simplified version of LSTM with two gates (reset and update), rendering it particularly beneficial for time series applications. GRU generally exhibits a similar level of efficiency and performance to LSTM (Pudikov & Brovko, 2020).

In this research, we demonstrated the application of AI methods to 3D cell culture images, which may enhance our understanding of the TME. Elucidation mechanisms on 3D TME may facilitate the development of customized treatments and personalized medicine (Walsh & Quail, 2023). The various cellular and acellular components constituting the TME may collectively influence tumor behaviors, including growth, invasion, metastasis, and treatment responsiveness (Elhanani et al., 2023). By combining advanced methods with AI and 3D multiplex platforms, we can make accurate predictions about how cells will evolve and behave over time. The 3D multiplex environments enable the simultaneous capture of a large amount of data, including biological replicates. Observing the cellular behavior and interactions within the TME is increasingly crucial for comprehending illness progression and elucidating disease mechanisms. We applied the DL transformer methodology to our image dataset and calculated raw data parameters.

The integration of advanced hybrid models enables us to understand cell behavior and elucidate the way for new therapeutic strategies. By leveraging the strengths of each hybrid ML model, scientists can gain deeper insights into the complex structure of GB progression and treatment responses. These multidisciplinary studies, which integrate *in silico* methods and cell culture techniques, can be valuable for patients and clinicians, while also enhancing the value of translational research, which refers to the process of converting scientific discoveries into clinical applications, thereby adding tangible value for patients by bridging the gap from bench to bedside. Future improvements in transformer-based DL methods may facilitate anomaly detection, medical image classification, and network architecture; however, additional investigation is required to obtain optimal results and increase accuracy.

2.6.4 Transformers

Considered an ML model, transformers have rapidly changed research perspectives in many domains, as they are highly performant at encoding complex sequences and function more accurately than other ML models, especially in very large databases. The following graphic (Figure 2.16) shows the data processing architecture for transformers (Lan et al., 2023). It demonstrates the transformer model, which has `N` encoder and `N` decoder units. Each encoder and decoder module has its own encoder and decoder blocks; in other words, there are `L` blocks in each module.

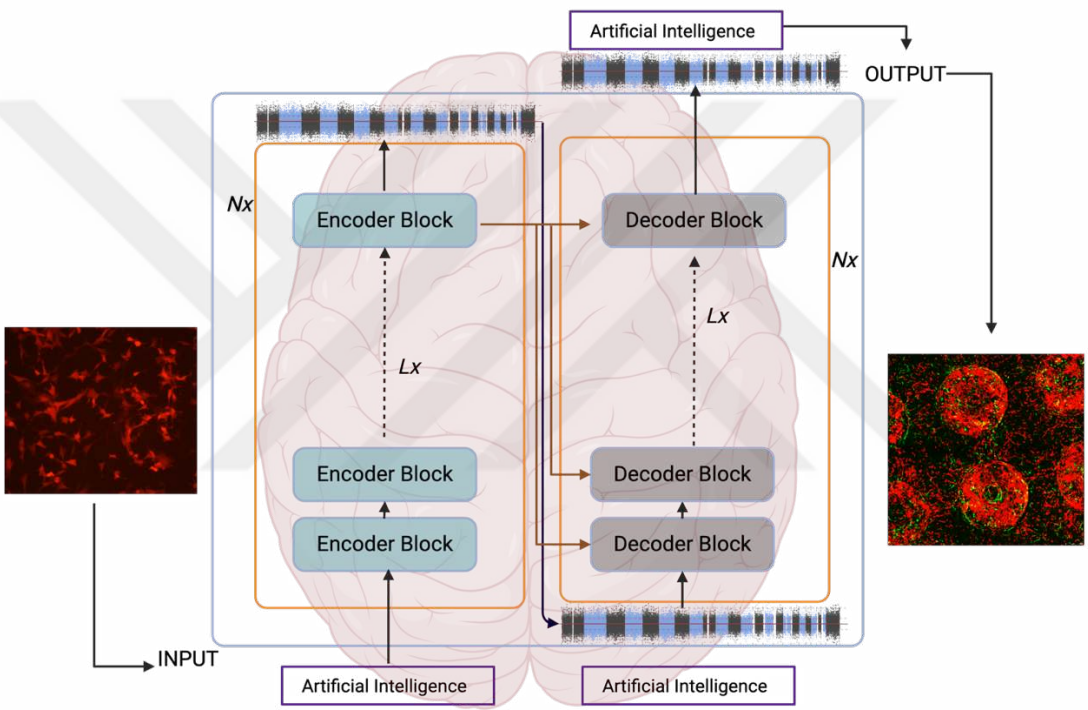

Figure 2.16: Transformer architecture explaining the encoder and decoder mechanisms based on the artificial intelligence approach for brain tumor images

In our study, we employed transformers as an ML model, specifically the Vision Transformer (`ViT`), due to the presence of images. To achieve this objective, we utilized the `timm` library, which integrates seamlessly with `PyTorch`.

The model received the calculated parameter results and the image dataset as input, enabling it to predict the migration capacity parameter on day 7 based on the results from days 3 and 5. The expected value was compared to the real value. The error

function was used to evaluate the accuracy of the ML model, and the results were then compared to those of previous studies that used separate ML models with the same input data.

2.6.5 CNN

CNN-based image classifiers have demonstrated high performance in manipulating large image datasets. In some studies, they played a role in distinguishing between diseased and healthy states, achieving accuracy rates of 80–100%. This success is attributed to high cell density, effective image preprocessing, and the use of deep CNNs with multiple hidden layers. Minimal image processing was required, with images divided into 64 tiles. Access to GPU resources facilitated the rapid processing of large datasets. Existing research reveals that combining classifiers based on tabular features with image data enhances explainability (D'Sa et al., 2023).

This study employed a hybrid machine learning model that integrates CNN architecture with LSTM, utilizing a sample dataset of 300 images characterized by time series characteristics.

2.6.6 Linear regression

Examining the function of mathematical models in analyzing spatio-temporal brain tumor progression, employing medical imaging modalities as time-series scans, is not a novel concept in the field of neuroscience. Some emphasize the transition from a 2D to a biomimetic 3D modeling framework that accounts for the spatial limitations of the peripheric areas. Scientists compare distinct nonlinear growth rates and cellular diffusion laws with actual scans of a GB patient's post-treatment results, delineating the effects of therapy on life durations and highlighting the implications of cell diffusion after surgical tumor excision (Tracqui & Mendjeli, 1999). Scientists try to understand the process of anticipating future occurrences, behaviors, and comprehending past events by identifying and defining the principles that govern them, primarily when deterministic equations are not known. They emphasize figuring out these rules from past data, utilizing examples like oscillatory systems. The concepts of understanding through mathematical insights and learning through algorithms, particularly neural

networks, are demonstrated to be two approaches to examining time series data (Weigend, 2013). Modern time series analysis encompasses three related topics. Statistical communication and control theory, the probabilistic theory of stochastic processes, and the statistical theory of regression and correlation analysis of time series data (Parzen, 1961). In our study, linear regression models are considered efficient when applied to small-sized raw datasets presented in a time-series format. The following table (

Table *2.2*) presents an example of wound-healing experimental results, which represent wound closure percentages for CTRL and TMZ-treated A127 and U87 GB cells, captured at 24 and 48 hours in both 2D and 3D imaging formats (Oğuz, 2025). Results 1, 2, and 3 represent the results of three biological replicates, which were captured simultaneously on ECM multiplex platforms. The predictions calculated in this table will be explained in detail in the following chapters (CTRL/TMZ Drug response comparison and predictions).

Table 2.2: Wound-healing experimental time series data of three biological replicates for CTRL and TMZ-treated human brain A127 and U87 GB cells (Oğuz, 2025)

SAAT	KONTROL/TMZ	Tümor Tipi	2D/3D	RESULT1	RESULT2	RESULT3
24H	CTRL	A127	2D.jpg	18,61	20,05	19,12
24H	TMZ	A127	2D.jpg	6,94	5,03	8,2
24H	CTRL	U87	2D.png	35,6	37,2	34,1
24H	TMZ	U87	2D.png	18,5	17,01	20,04
24H	TMZ	U87	3D.jpg	3,8	3,3	4,5
24H	CTRL	A127	3D.png	9,2	8,9	10,4
24H	TMZ	A127	3D.png	4,7	5,3	5,8
24H	CTRL	U87	3D.png	11,5	13,9	12,34
48H	CTRL	A127	2D.png	38	40,98	43,23
48H	TMZ	A127	2D.png	17,25	15,12	19,54
48H	CTRL	U87	2D.png	48,5	50,3	45,6
48H	TMZ	U87	2D.png	32,7	30,09	28,5
48H	CTRL	A127	3D.jpg	26,2	23,4	27,6
48H	CTRL	U87	3D.jpg	47,9	45,4	40,8
48H	TMZ	U87	3D.jpg	6	7,9	5,1
48H	TMZ	A127	3D.png	5,5	4,8	7,5

The 95% confidence intervals were computed using the RMSE and a sample size of $N = 72$, resulting in an interval width of $\pm 0.234 \times RMSE$. MTT reduction was

measured via a plate reader spectrophotometer at 540 nm, with 650 nm serving as the reference wavelength.

The following equation explains Confidence Intervals (CI) for based on the reported RMSE value.

$$\mathrm{CI}_{95\%} = \bar{y} \pm t_{0.975,\,N-1} \cdot \frac{\mathrm{RMSE}}{\sqrt{N}} \tag{2.5}$$

Cell viability was determined utilizing the following formula where A_{540} indicates the absorbance at 540 nm, which is the peak wavelength for colorimetric cell viability tests like MTT or XTT. A_{650} indicates the absorbance at 650 nm, which is used as a reference to correct for nonspecific absorbance and background noise.

$$\text{Cell viability (\%)} = \frac{(A_{540}^{\text{sample}} - A_{650}^{\text{sample}})}{(A_{540}^{\text{control}} - A_{650}^{\text{control}})} \times 100 \tag{2.6}$$

The table below (Table 2.3) presents time-series data on MTT viability from three separate biological replicates of U87 GB cells grown in 3D, both under control conditions and after treatment with TMZ at varying doses. Each row represents a particular experimental condition characterized by TMZ concentration and incubation duration. At the same time, the columns display the individual MTT viability measurements for the three replicates together with their mean value. The structured and nearly monotonic temporal trends in these raw time-series data suggest that linear regression models are suitable for elucidating the underlying time-dependent behavior of cell viability, especially for short- to mid-term forecasts, while also facilitating quantitative assessment of dose-dependent cytotoxic effects and experimental reproducibility. Untreated cells were deemed to reflect 100% viability. The real and predictive results will be explained in detail in the next chapter (Results, CTRL/TMZ Drug response comparison and predictions).

Table 2.3: MTT viability experimental time series data of three biological replicates for CTRL and TMZ-treated U87 GB cells in 3D for various TMZ doses (Oğuz, 2025)

SAAT	KONTROL/TMZ	Tümor Tipi	2D/3D	TMZ DOSE	RESULT1	RESULT2	RESULT3
24	CTRL	U87	3D	0	99,511	100,049	100,538
24	TMZ	U87	3D	10	94,377	95,257	95,159
24	TMZ	U87	3D	25	88,606	87,237	89,34
24	TMZ	U87	3D	50	87,433	84,303	87,042
24	TMZ	U87	3D	100	80,293	79,315	78,973
24	TMZ	U87	3D	250	67,042	67,579	67,971
48	CTRL	U87	3D	0	99,159	100,374	100,514
48	TMZ	U87	3D	10	92,991	83,645	95,794
48	TMZ	U87	3D	25	81,682	81,729	82,009
48	TMZ	U87	3D	50	74,766	81,963	74,533
48	TMZ	U87	3D	100	71,355	70,981	71,215
48	TMZ	U87	3D	250	57,336	57,056	57,383
72	CTRL	U87	3D	0	99,459	98,79	101,803
72	TMZ	U87	3D	10	74,478	73,294	75,843
72	TMZ	U87	3D	25	67,937	67,963	63,405
72	TMZ	U87	3D	50	67,937	67,963	63,405
72	TMZ	U87	3D	100	49,678	60,237	62,941
72	TMZ	U87	3D	250	39,042	44,399	47,257

CHAPTER 3: RESULTS

3.1 Examining GB cell behaviors with parameters

This study examines the migration capacity and the proliferation rate of U87 human GB cells. The following figure (Figure 3.1) illustrates the growth of RFP-tagged U87 cells and GFP-tagged human astrocytes over a nine-day culture period. From day 3 to day 9, there is a significant increase in the number of cells, demonstrating the rapid growth of GB cells. Nonetheless, a slight decrease in proliferation is observed between days 7 and 9. This decrease may be attributed to cellular apoptosis or contact inhibition, a phenomenon in which cells overlap and limit further proliferation due to spatial constraints and resource competition. To resolve this issue, we may calculate volume rather than area. From a biological perspective, these findings are consistent with the documented aggressiveness of U87 cells, which are known for their rapid proliferation and adaptability. The observed reduction at later time points may further indicate microenvironmental constraints, such as those found in solid tumors. This activity demonstrates how malignant U87 cells can be and provides information about how tumors grow, as well as the challenges that arise when attempting to treat GB growth.

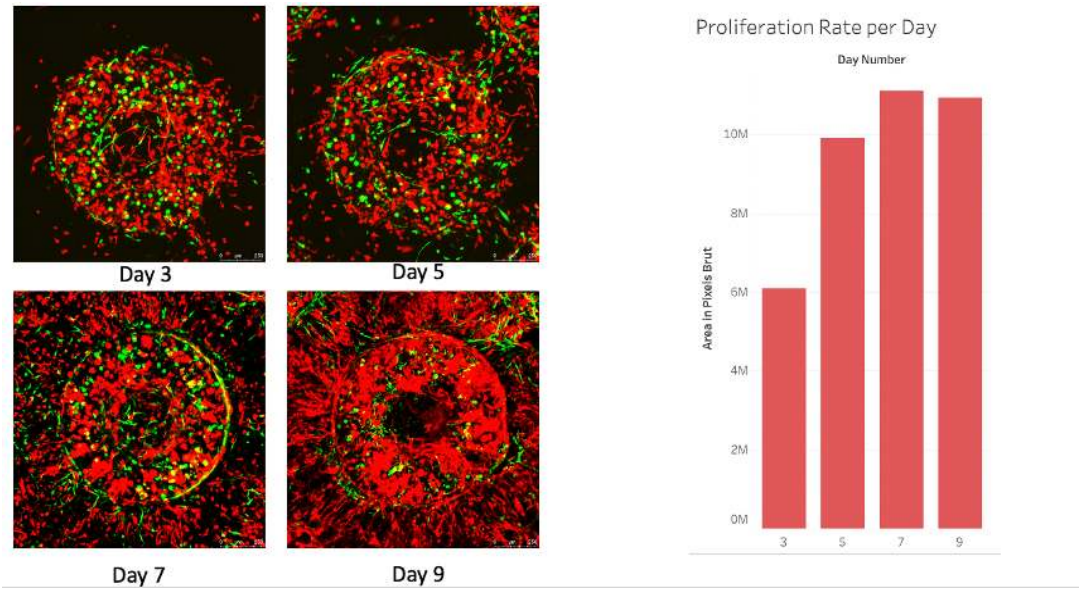

Figure 3.1: Migration capacity and proliferation behavior of RFP-tagged human brain U87 human GB cells (Ismayilzada, 2024)

The graph of the results in the following figure (Figure 3.2) demonstrates how the migration capacity of human brain U87 GB cells evolves on separate days. The X-axis represents the days of the experiment, as explained in Section 2.2, and the Y-axis displays the calculated migration capacity parameter and the aggregated results, averaged per day. We observe a clear trend of significant improvement in migration capacity, with a steady increase from day 3 to day 9. This behavior demonstrates how U87 cells behave, showing that their ability to move increases with longer culture periods.

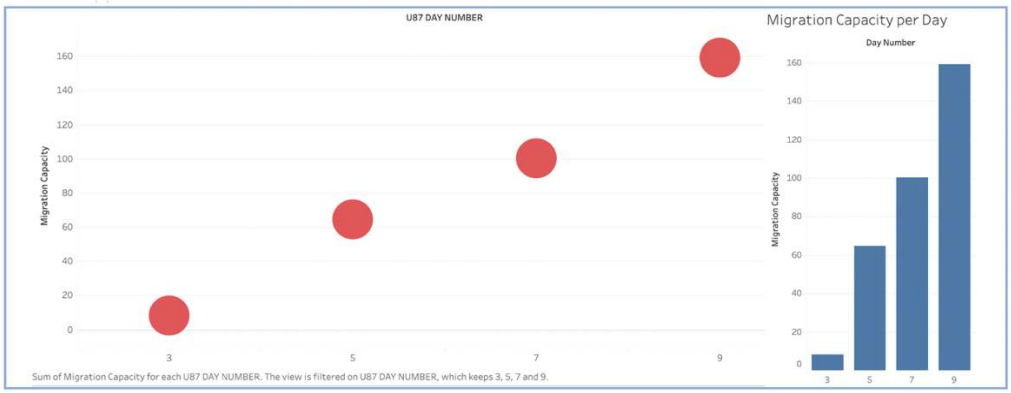

Figure 3.2: Results graph, migration capacity parameter, per day

Another scientific representation is shown in the following figure, which demonstrates how the migration capacity changed over nine days under two distinct culture configurations: direct co-culture and stamping co-culture (Figure 3.3). The details of these configurations can be found in previous chapters (Figure 2.3: Migration behavior of U87 GB cells, comparison of direct and stamping configurations for migration capacity). The left line graph illustrates migration capacity at four separate times: Day 3, Day 5, Day 7, and Day 9. Cells cultivated in direct co-culture demonstrate a significant and incremental enhancement in migration capacity, increasing nearly linearly from Day 3 to Day 9. The stamping configuration, on the other hand, remains almost flat every day, indicating little migration behavior. The significant distinction between the two circumstances indicates that direct co-culture plays a substantial role in driving cell migration, whereas stamping configuration alone has a minimal effect. The right bar graph illustrates a statistical comparison of migration capacity in the direct co-culture configuration over distinct time points. From Day 3 onward, there is a significant increase in the migration capacity at each time point. The numerous substantial changes ($p<0.001$) seen over days further corroborate the reliability of this upward pattern.

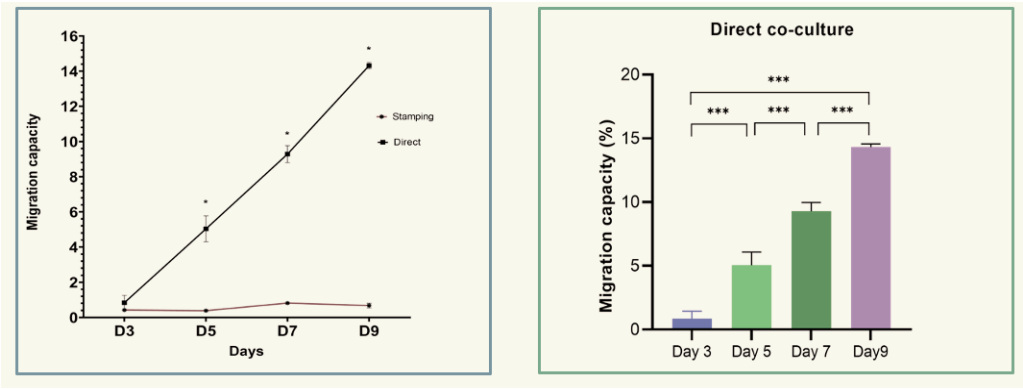

Figure 3.3: Results of the migration capacity parameter of U87 GB cells over days

The ML models used to predict migration capacity on day 7 based on data from days 3 and 5 performed accurately, as shown in the table above. The values, all exceeding `0.80`, demonstrated that the training performed accurately and that the models can capture significant temporal patterns in the dataset effectively. By changing the density number parameter in Python from 10 to 1, we achieved better accuracy, with a value of `0.85`. This highlights the importance of adjusting hyperparameters to achieve the most

accurate results for an ML model. This finding suggests that a systematic analysis of model parameters could significantly improve predictability. In addition, expanding this research to include more hyperparameters, varied neural network topologies, and the implementation of regularization techniques may enhance the understanding of optimizing prediction reliability. These findings underscore the importance of precise calibration and methodological refinement in advancing predictive modeling for cellular migration behaviors.

3.2 Transformers prediction results

We used the same input data from days 3 and 5 to train ML models to predict the day-7 migration capacity parameter. The transformer-based architecture produced an RMSE of `0.84`, a R^2 score of `0.1791`, an MSE of `0.7167`, and an MAE of `0.5954`. The MAE value of about `0.60` shows that the average distinction between predicted and actual values is minimal in absolute terms. However, the low R^2 score indicates that the model explains only a small portion of the variance in the migration capacity outcome. As a result, the transformer model does not accurately capture the underlying prediction patterns, particularly when performance measurements are still below the estimated threshold of 80%.

The following graph (Figure 3.4), which compares actual and expected migration capacity values, demonstrates this restriction even more clearly: errors are significant in relative terms, with projections often being away by over a factor of two from the genuine values. This pattern indicates possible limitations in the data and modeling methodology. The dataset may not be large enough to assist strong generalization, and the current set of features may not include all considerable predictors. Adding more variables to the dataset, including culture conditions, experimental metadata, or spatial and morphological cell features, may improve the model's accuracy for transformers. The RMSE (0.8466) was used to find the standard deviation of the residuals, which was used to estimate prediction uncertainty. Therefore, a 95% prediction interval (PI) was found to be ±1.66 log units around each predicted value. The model's low explanatory power ($R^2 = 0.1791$) and the fact that some samples have considerable deviations suggest a high level of uncertainty in the predictions. The 95% prediction interval (PI)

was approximated by assuming normally distributed residuals and was computed using the following equation for approximate 95% prediction interval (PI) for model predictions.

$$PI_{95}\% = \hat{y}_i \pm z_{0.975} \cdot \sigma_\varepsilon \quad (3.1)$$

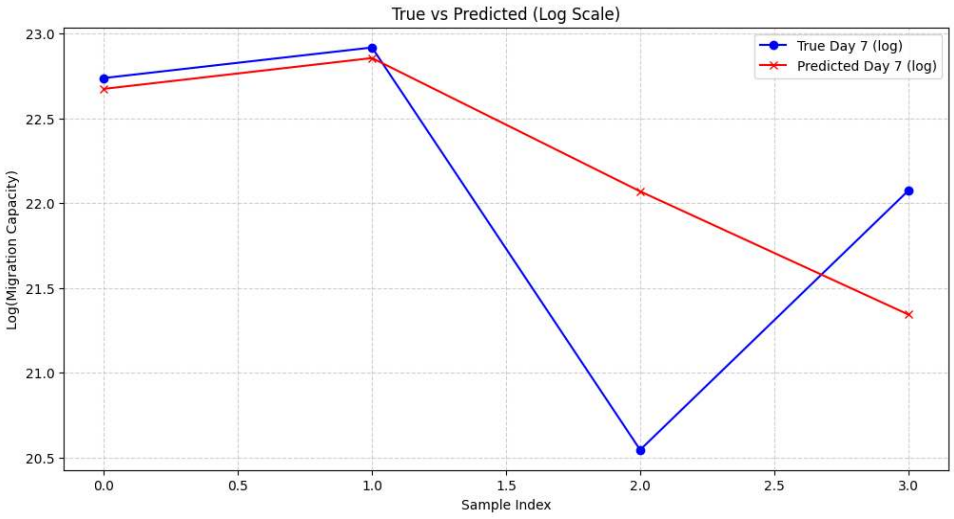

Figure 3.4: Prediction results with the transformers ML model, R^2 Score: 0.1791, MSE: 0.7167, RMSE: 0.8466, MAE: 0.5954, PI: ±1.66

3.3 Hybrid test prediction results

The four-panel graphic in the following figure (Figure 3.5) illustrates the performance of the RNN–LSTM model on training and test datasets according to the pre-activation rate. Each subplot presents the projected pixel values alongside the actual measurements, facilitating straightforward evaluation of the model's accuracy and stability. The test-set results (top row) indicate that the model accurately captures the overall trend of the real data for both the Density32 and Density16 groups. The projected curves closely follow the changes in the observed data, indicating that the model has effectively learned the global temporal patterns. Still, there are apparent disparities at locations where projections are far lower than the actual values. These outliers may indicate that the model is sensitive to sudden changes in the area, that the

training set lacks sufficient examples of specific patterns, or that the real measurements are noisy. The model exhibits remarkable fidelity in the training set results (bottom row), with predicted values that closely match the actual measurements at nearly all sample locations. The close grouping of prediction, ground truth pairs shows that the model is well-fitted and has accurate internal consistency. However, vertical drops that occur from time to time indicate that some individual samples remain challenging for the network to predict, even after being shown to it many times during training. These distinctions could be due to biological variability, experimental noise, or the fact that the model cannot capture infrequent or non-linear changes. The graphic (Figure 3.5) illustrates that the RNN–LSTM model performs well in various density situations. It aligns well with real values and can be used for both training and testing sets. Localized prediction errors indicate that the model still requires further improvement. This could involve adding more features, enhancing data preprocessing, or employing regularization techniques to ensure the model remains robust when presented with irregular data points.

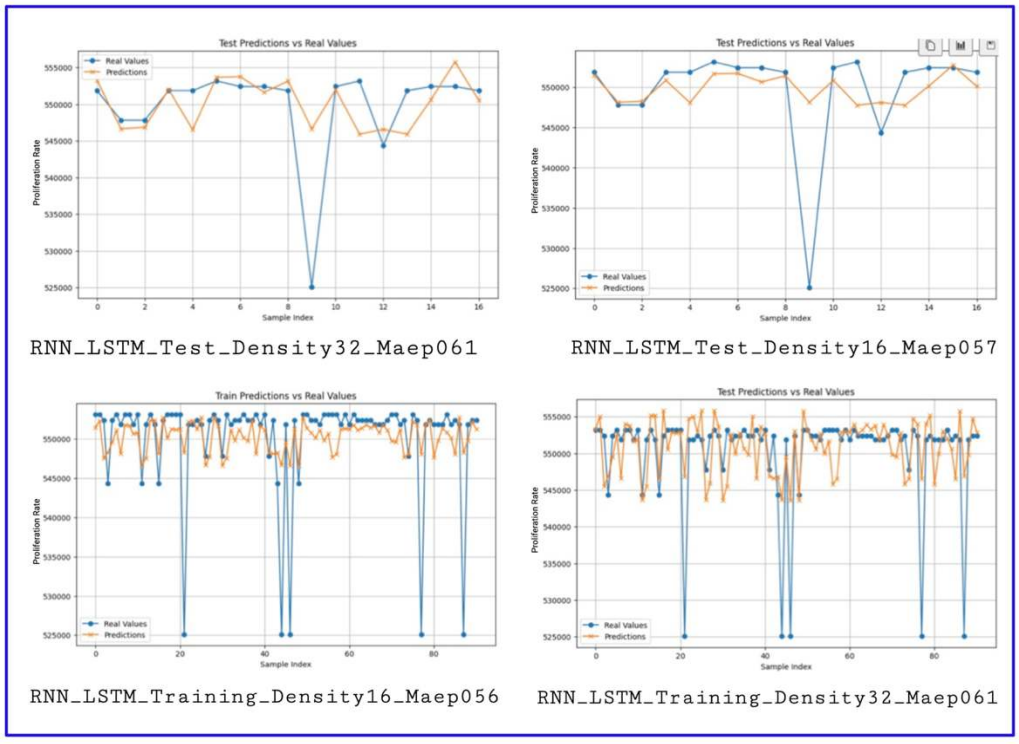

Figure 3.5: Hybrid ML model prediction results RNN/LSTM for proliferation rate parameter

The four panels in the following figure (Figure 3.6) present a comparative evaluation of hybrid deep learning architectures, CNN-LSTM and RNN-GRU, applied to training and test datasets under distinct density conditions (85% training and 15% test dataset). Each graph has a sample index representing data points and proliferation rate, in terms of $pixel^2$.

Model predictions are represented by the orange line, and real values are represented in blue, allowing for a direct assessment of predictive accuracy and temporal stability. In the CNN-LSTM test results (Density 16), the model demonstrated moderate alignment with real values, although several fluctuations and pronounced deviations are evident across sample indices. The variability in prediction accuracy is reflected in the relatively higher MAE value (MAE ~ 292), indicating that while the model captures the general trend, it struggles with finer-scale variations in the dataset.

On the other hand, the RNN-GRU test performance (Density 32) shows noticeably improved alignment between predicted and actual values. The prediction curve exhibits reduced amplitude fluctuations and fewer extreme deviations relative to the CNN-LSTM test output. This more substantial conformity is supported by a lower error metric (MAE ~ 79), suggesting that the GRU-based hybrid architecture is better suited for modeling the temporal dependencies present in the test data. The training curves illustrate similar patterns.

The CNN-LSTM training plot (Density 16) reveals substantial divergence between predicted and real values across many sample points, with irregular prediction spikes contributing to a comparatively high training error (MAE ~ 250). In contrast, the RNN-GRU training curve (Density 32) exhibits tighter clustering of predictions around the real data line, resulting in a notably low training error (MAE ~ 0.78). The smoother and more consistent predictive behavior indicates that the GRU-based model achieves a more stable learning process and improved generalization.

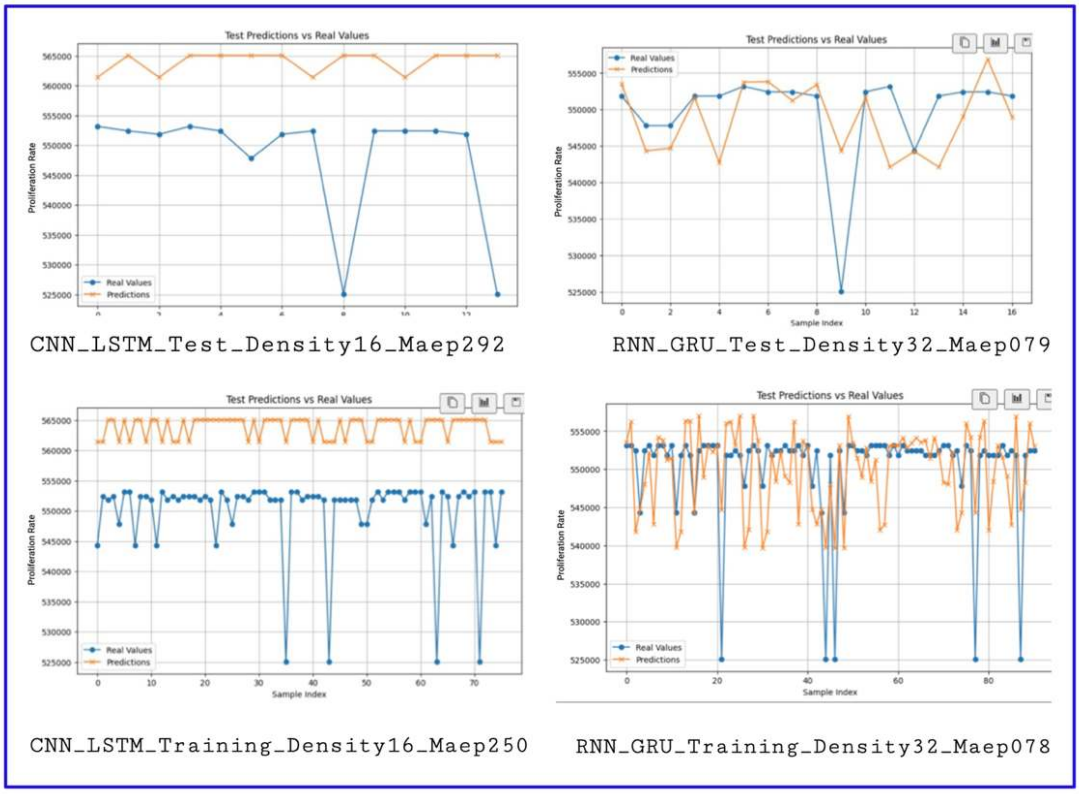

Figure 3.6: Hybrid ML model prediction results, CNN/LSTM & RNN/GRU for proliferation rate parameter

The CNN-LSTM model achieves lower MAEP values on the training set (MAEP = 250) compared to the test set (MAEP = 292). This indicates a moderate generalization gap. On the other hand, RNN-GRU models have error distributions that are more stable across training and test datasets. The following table (Table 3.1) compares the ML models in terms of performance. The GRU architecture had the highest overall accuracy (0.83), while the LSTM architecture had the lowest MAE for both migration capacity (0.158) and proliferation rate (0.061). For the CNN/LSTM model, with three hidden layers, we observed a higher MAE for predicting proliferation rates (0.292) compared to the other models. Prediction intervals were approximated as ±1.96 × MAE, assuming normally distributed residuals.

Table 3.1: Performance comparison of CNN/LSTM and recurrent models, including overall accuracy, mean absolute error (MAE), and approximate 95% prediction intervals (PI) for migration capacity and proliferation rate

Model	Overall Accuracy	Migration Capacity MAE	Migration Capacity PI (±)	Proliferation Rate MAE	Proliferation Rate PI (±)
CNN-LSTM	N/A	N/A	N/A	0.292	0.572
RNN	0.80	0.225	0.441	0.079	0.155
LSTM	0.81	0.158	0.310	0.061	0.120
GRU	0.83	0.163	0.319	0.078	0.153

The same information with logarithmic migration capacity results is demonstrated in the following graph. To facilitate a direct and comparative analysis of the configuration between experimentally measured and model-predicted migration capacities, both datasets underwent logarithmic transformation and normalization using a robust scaling method. This approach ensured that minor temporal variations could be evaluated independently of absolute magnitude distinctions. To ensure consistent scaling between experimental and model-derived data, both actual and projected migration capacity values underwent a base-10 logarithmic transformation. This step minimized the effect of disparities in absolute magnitude, facilitating the comparison of relative variations within the constrained dynamic range (approximately 5.25 × 105 to 5.55 × 105). A strict scaling method was employed, with the pooled median serving as the primary reference point. The interquartile range (IQR) was used to normalize the data and lessen the effect of any outliers. The adjusted values were then confined within a ±1 IQR interval to mitigate extreme fluctuations and were linearly scaled to the [0, 1] range. This combination of logarithmic and robust normalization preserved the natural patterns in both datasets while equalizing their amplitudes, which facilitated the comparison of the model's visual and quantitative predictions of migration behaviors that had been empirically validated.

The following figure (Figure 3.7) demonstrates the results of the RNN-LSTM hybrid model on the test dataset. The x-axis refers to individual data points from image

parameter calculations, and the y-axis refers to the normalized migration capacity values. Both real and predicted migration capacity parameters have been logarithmically normalized and robustly scaled, enabling a comparison of results across a wide range of measurements. These preprocessing processes reduce the effect of outliers and irregular distributions; thus, the model's outputs are more interpretable. The results displayed are from the RNN–LSTM hybrid model on the test dataset. The near-perfect match between predicted and real values for most data points indicates that the model accurately reflects the underlying patterns of migration behavior.

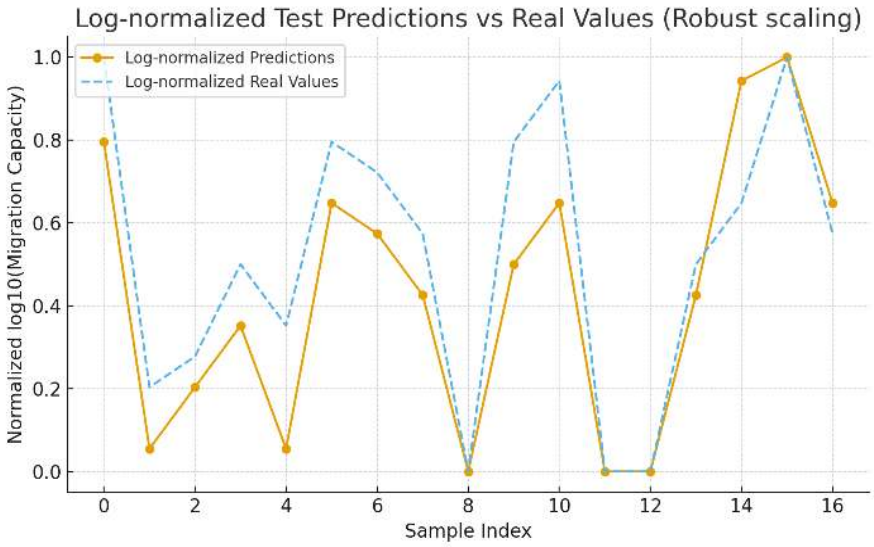

Figure 3.7: Logarithmic normalization and robust scaling of real and predicted migration capacity values.

The accuracy measures how many of the anticipated outcomes were correctly predicted, indicating how accurately the model performs. The ML models used in this research to predict day-7 proliferation rate based on day-3 and day-5 observations displayed significant accuracy. The RNN model achieved an accuracy of 0.80, the LSTM model achieved an accuracy of 0.81, and the GRU model achieved an accuracy of 0.83. All these values exceed the usual cutoff of 0.80, indicating that the models accurately captured the underlying temporal patterns and provided reliable predictions. The GRU model also improved when the density parameter in the Python code was changed from

10 to 1, resulting in an accuracy of 0.85. This increase suggests that the model's sensitivity to input density affects its predictive accuracy, and that fine-tuning these hyperparameters can lead to significant improvements in accuracy. In general, the results demonstrate that RNN architectures can be effective in predicting cell behavior, with GRU yielding the most accurate results in this specific scenario.

3.4 Voxel-based volume calculation on 3D confocal images

Using this computational approach, the volume of the 3D confocal image of RFP-tagged U87 glioblastoma cells at day 14 (Figure 2.12: 3D confocal image of U87 and BMEC cells in stamping configuration, day 14, scale 200 µm) was calculated to be 729312 μm^3. The following figure (Figure 3.8) shows separate slices, including vertical, horizontal, and cross-sectional perspectives of the same image.

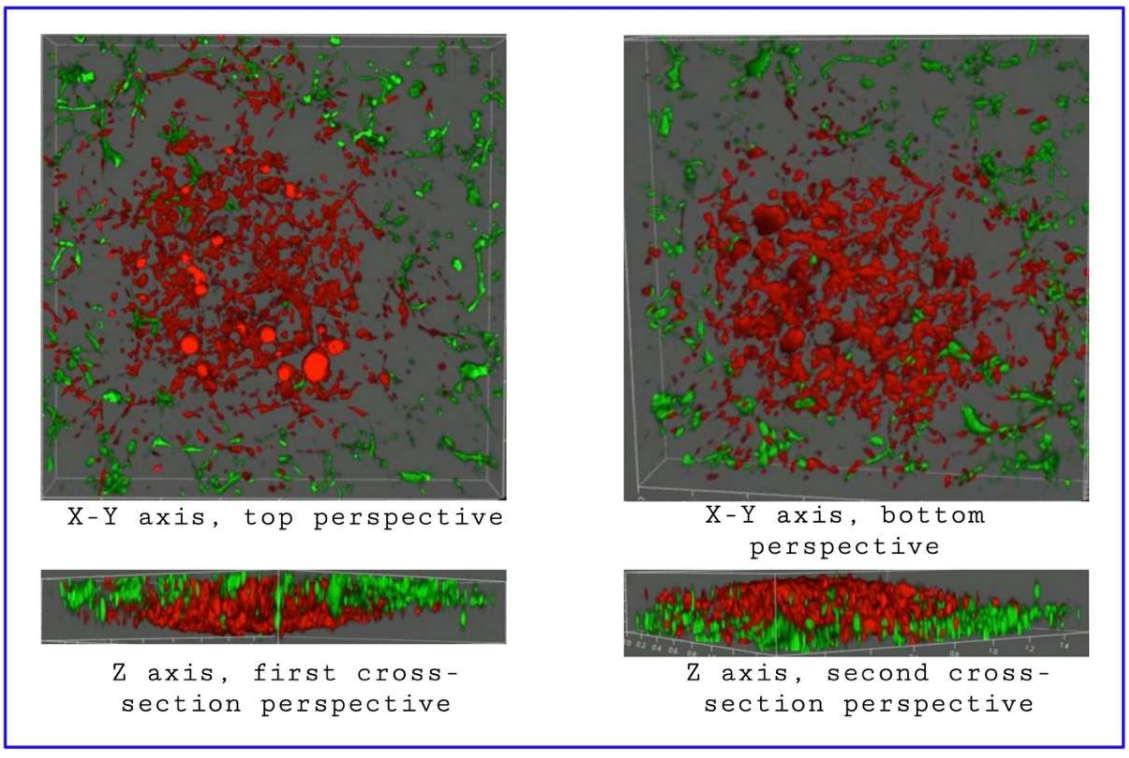

Figure 3.8: Vertical, horizontal, and cross-section slices of the 3D confocal image of U87 and BMEC cells in stamping configuration, day 14, scale 200 µm

3.5 Live cell imaging results

In the context of this thesis, live cell imaging techniques can provide valuable insights into understanding GB cell behaviors, such as migration and proliferation. Details concerning methodologies are presented in previous chapters (Live cell imaging). The following figure (Figure 3.9) illustrates all the frames of the RFP-tagged U87 2D live cell imaging video in a mosaic. This video was recorded over approximately 66 hours, equivalent to 3 days; each frame represents a photo captured by the Zeiss confocal microscopy at the end of each hour.

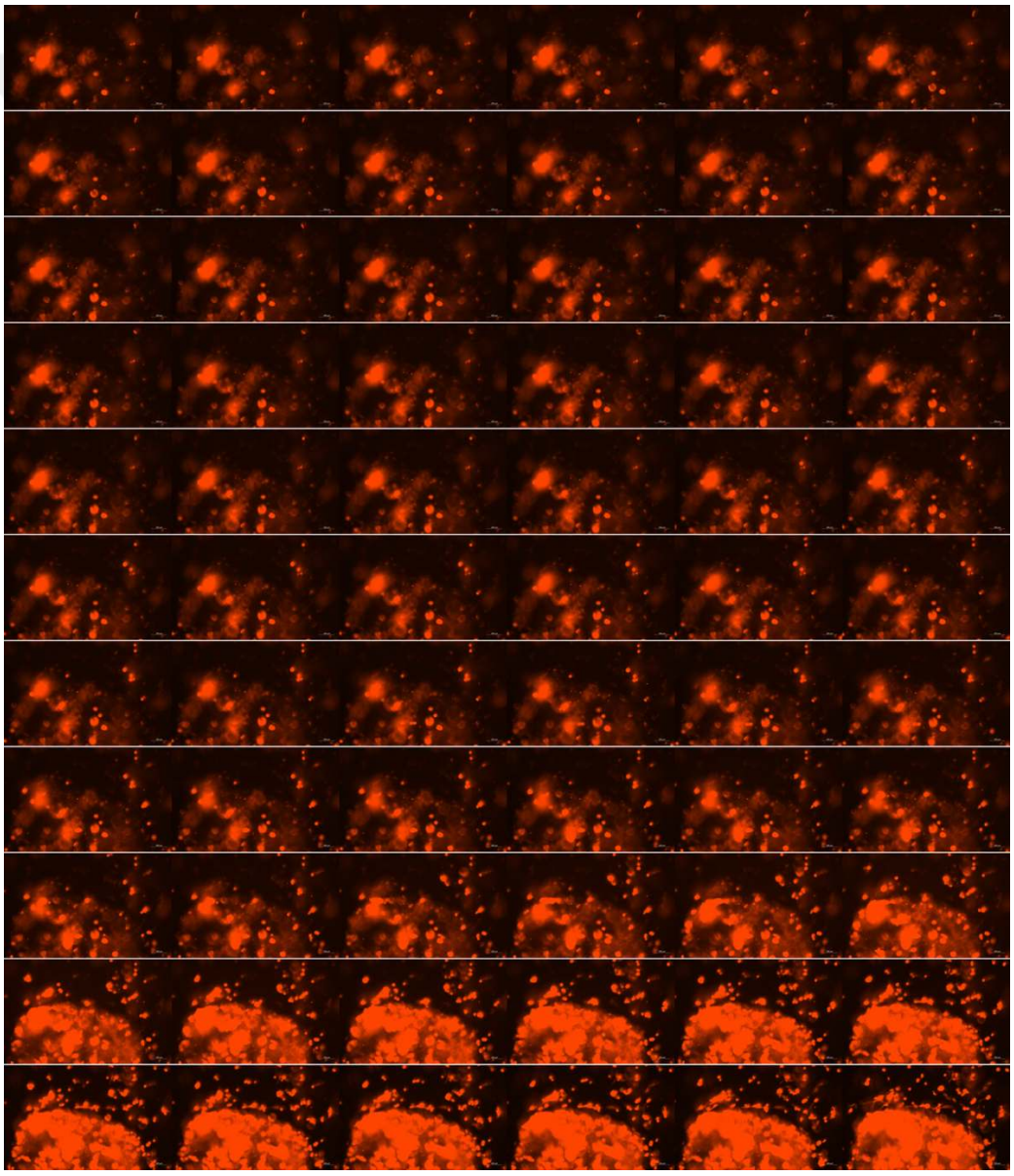

Figure 3.9: Frames of live cell imaging video captured during 3 days

The following figure (Figure 3.10) presents the temporal dynamics of RFP-tagged U87 cells, as captured using live-cell imaging techniques. The x-axis represents the frame index, indicating the hour of image acquisition, while the y-axis measures the proportion of the imaged area occupied by RFP-tagged U87 cells. The continuous blue line indicates cell-occupied areas, and the dotted orange line denotes acellular medium, allowing for clear differentiation and analysis of cell proliferation over time.

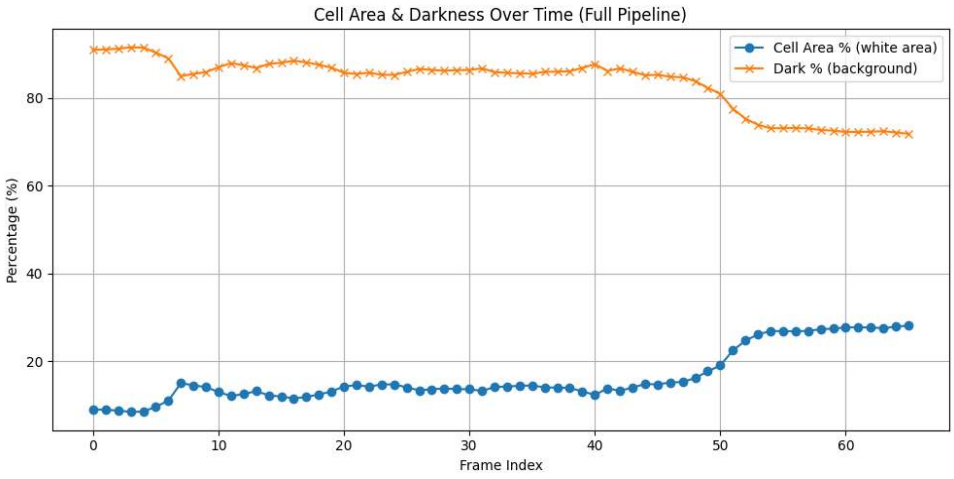

Figure 3.10: Live cell imaging temporal variation results for RFP-tagged U87 cell coverage in percentage

The following figure (Figure 3.11) depicts the temporal progression of RFP-tagged U87 cell expansion as quantified through image-based area measurements. The x-axis represents the frame index, which corresponds to the specific hour at which each image was acquired, enabling a longitudinal assessment of cellular growth or migratory behavior. The y-axis denotes the total pixel area occupied by the fluorescently labeled U87 cells, providing a direct quantitative indicator of changes in cell coverage over time. The use of a continuous blue line to represent cell-occupied area ensures clear visualization of temporal trends, highlighting fluctuations or progressive increases that may reflect underlying biological processes such as proliferation, migration, or morphological alterations. This representation offers a precise and interpretable measure of cellular dynamics across the imaging period.

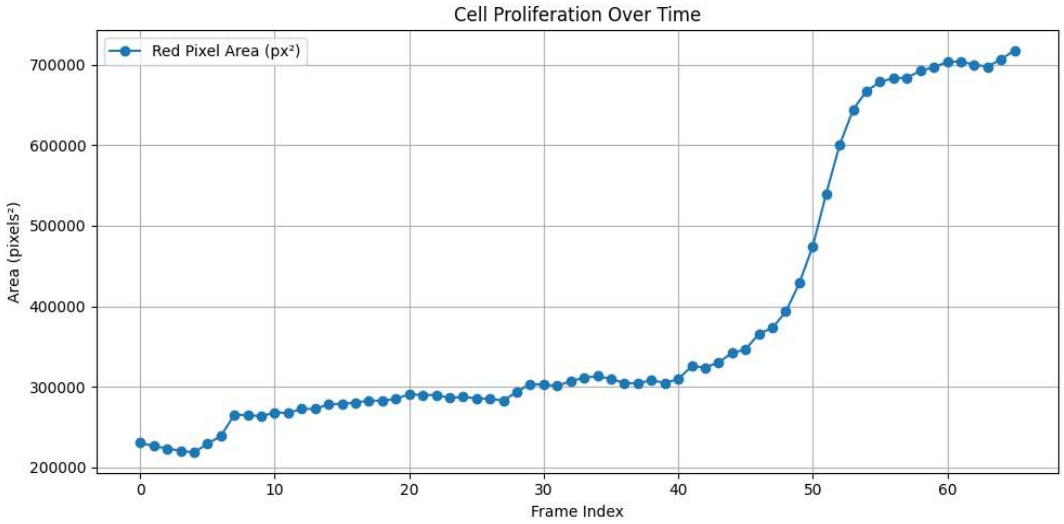

Figure 3.11: Live cell imaging of temporal changes in RFP-tagged U87 GB cell growth

3.6 CTRL/TMZ Drug response comparison and predictions

The following graph (Figure 3.12) illustrates a comparison of how cells respond to CTRL and TMZ treatment at three distinct time points: 24 hours, 48 hours, and 72 hours. In this context, the 24 hours and 48 hours represent real values, while the 72nd hour is predicted by ML algorithms. The results are displayed for three averaged measurements (Avg. Result of 3 biological replicates, Result 1, Result 2, and Result 3 of

Table *2.2*, Linear regression), facilitating the assessment of temporal patterns and treatment-specific effects. The control group demonstrates a consistent and significant increase from 24 hours to 72 hours. This means that untreated cells are growing or becoming more active in a usual way. The TMZ-treated group, on the other hand, consistently has far lower numbers, which demonstrates that TMZ prevents cells from functioning or staying alive. The inequality between CTRL and TMZ becomes more pronounced over time, indicating that TMZ effects intensify with prolonged exposure. This indicates that early measures, 24 h and 48 h, can be used to predict later behavior, 72 h reliably. In other words, the predictable trajectory indicates that initial medication responses may function as reliable predictors for modeling long-term therapy effects in ML models.

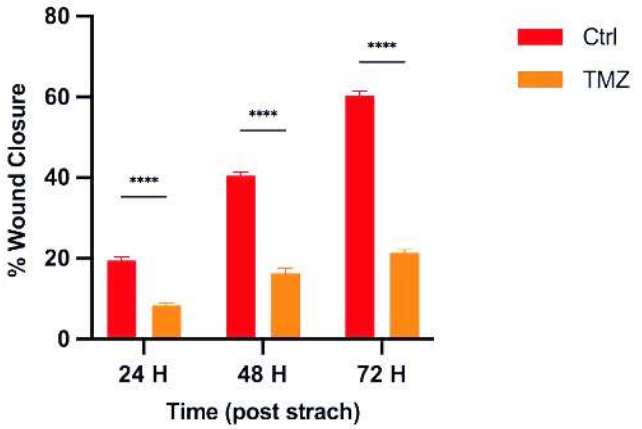

Figure 3.12: Prediction of 72nd hour for CTRL & TMZ human GB cells drug responses

In the following figure (Figure 3.13), the four panels illustrate how TMZ affects cells over time at varying doses (0–250 mg) in A172 and U87 GB cell lines grown in 2D and 3D systems. In all of the graphs, TMZ 0 mg is the control group that was not treated with the drug, and higher dosages show more TMZ exposure. The solid lines represent values measured in experiments (24, 48, and 72 hours), whereas the dotted lines indicate predicted values at later times (96, 120, and 144 hours) by regression-based ML algorithms. In both human brain cell lines (A127 and U87, both GB cells) and culture formats (2D and 3D), the control groups consistently exhibit high viability values, exhibiting only a minor decrease over time. This shows that the baseline growth is consistent in the CTRL group. All groups treated with TMZ, on the other hand, exhibit dose- and time-dependent decreases in response values, indicating that TMZ has the expected cytotoxic effect. The two types of GB cell lines, *Homo sapiens* brain A172 and U87 GB models, experience minor to moderate declines (10–50 mg) at lower dosages, whereas higher doses (100–250 mg) lead to a more pronounced early drop, which can be observed as early as 24 hours. These decreases are consistently observable in the 3D cultures, suggesting that 3D systems may exhibit enhanced drug sensitivity or more physiologically relevant drug penetration dynamics in comparison to 2D cell culture studies. The projected values (dotted lines) are based on observed changes over time. In every case, expectations follow the path of the values found in the experiment. Control forecasts remain unchanged, which aligns with the slight drop observed in the real data. If predictions yield negative results, the expected numbers are set to zero, as viability cannot be a negative value. TMZ-treated predictions continue to decline,

indicating the strength of the dose-dependent inhibitory pattern. Predictions for high doses approach complete loss of viability, which aligns with the early significant decreases observed in the experiment data. The prediction curves must remain parallel to the real data, demonstrating that the model accurately indicates how the drug's effects change over time with various dose levels, even after the times recorded in the experiment. This consistency means that early responses (24–72 hours) can be used to predict later behavior (96–144 hours) with a fair amount of accuracy with ML algorithms. These results demonstrate that TMZ consistently and significantly reduces cell viability in both the A172 and U87 cell lines. The effects are stronger at higher dosages and in 3D cultures. The correlation between observed and predicted results underscores the potential for early-timepoint data to enhance long-term medication response modeling in GB research.

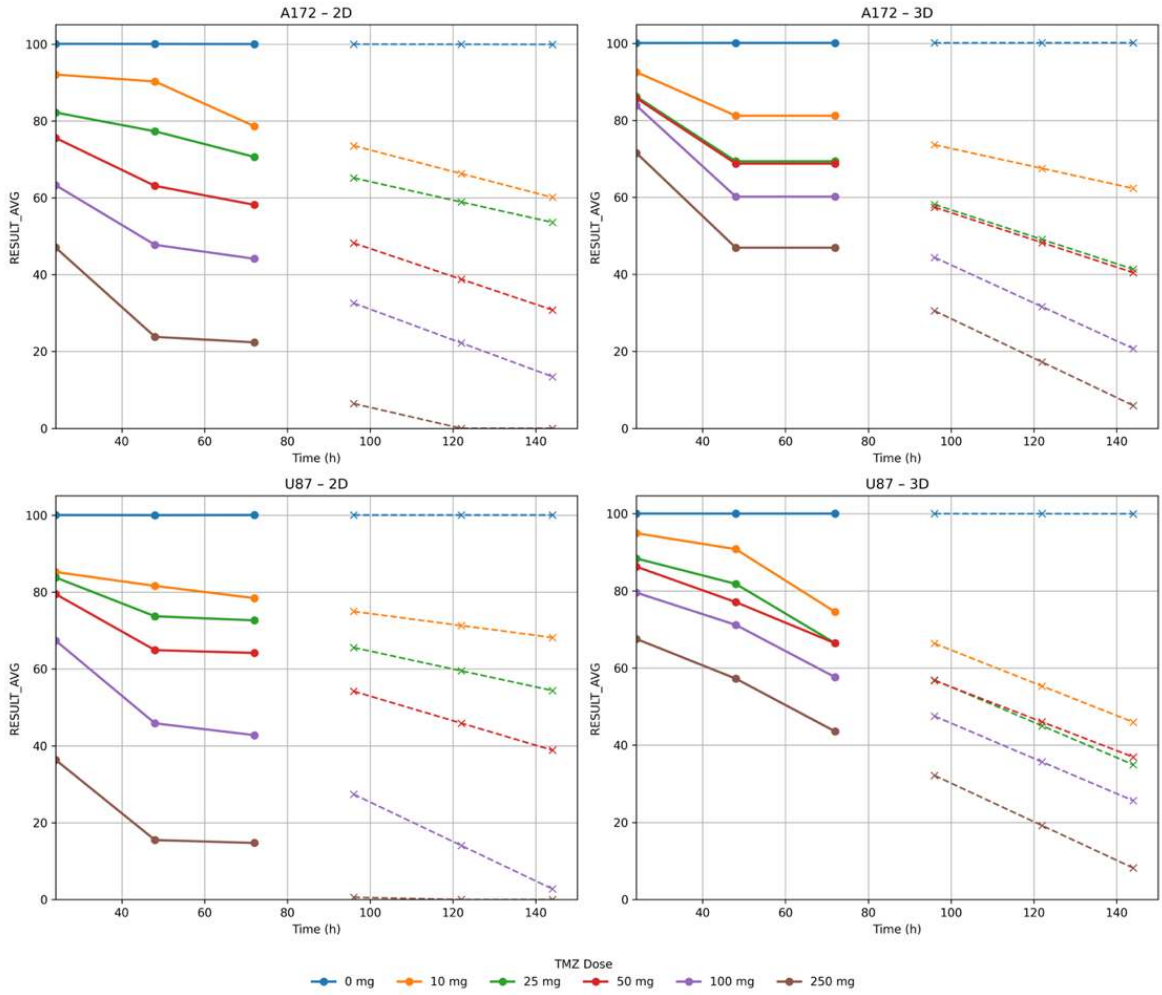

Figure 3.13: TMZ dose–response real values and predictions in 2D & 3D A172 and U87 Cells

The following graphs (Figure 3.14) illustrate the actual and predicted 72-hour MTT viability values derived from linear regression models, utilizing 24-hour and 48-hour observations, for human brain A172 and U87 cell lines cultivated under both 2D and 3D configurations. Initially, in the context of this experiment, experimental results at 24 hours, 48 hours, and 72 hours exist as absolute values. Each panel represents a distinct experimental group, with the solid diagonal line indicating complete concordance between predicted and actual results.

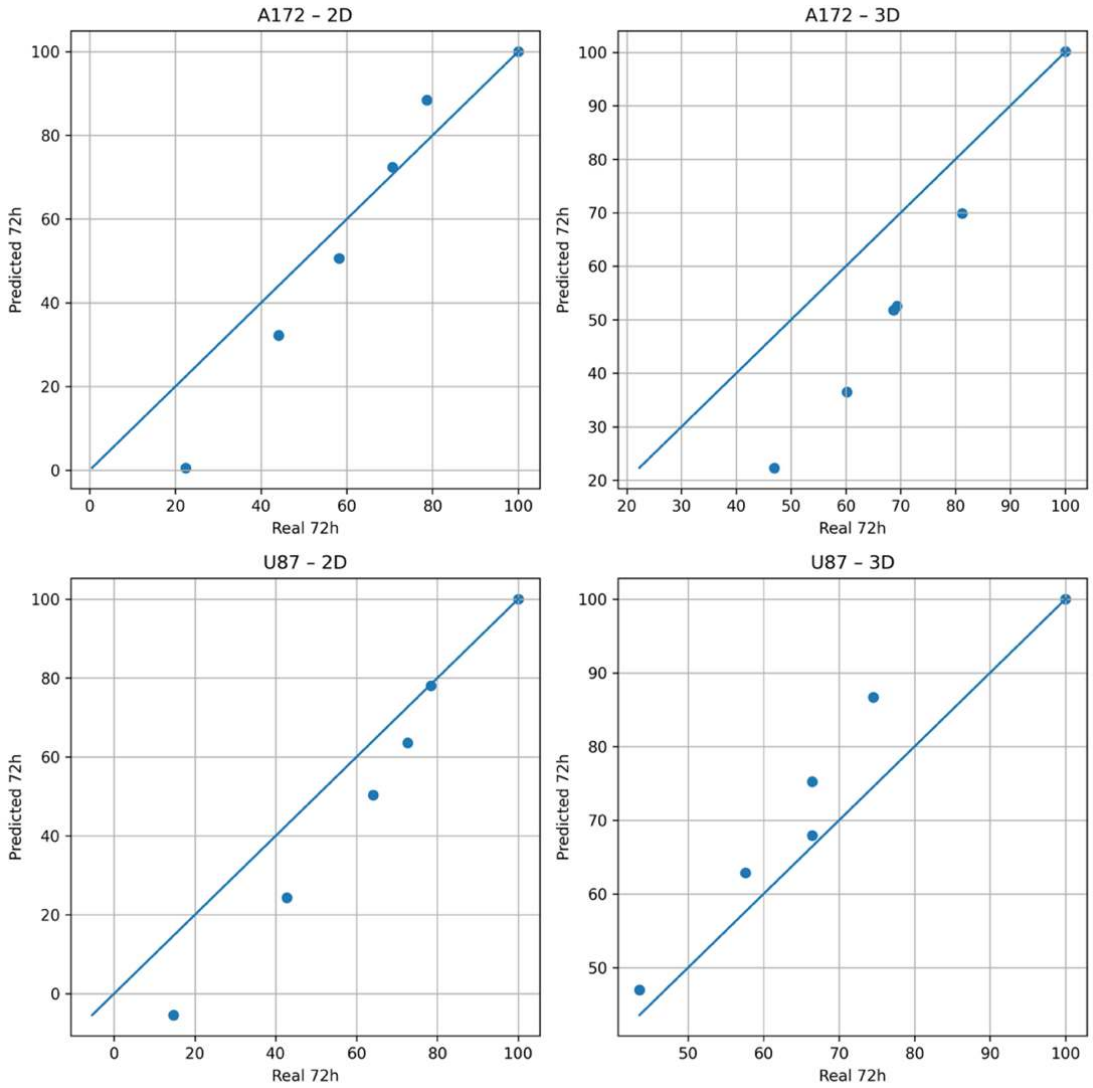

Figure 3.14: Comparison of real and predicted 72-h cell viability results for human brain A127 and human brain U87 cells in 2D and 3D ECM platforms

Linear regression models were trained using only 24 h and 48 h MTT measurements and used to predict 72 h viability. Model performance was evaluated by comparing predicted and experimentally measured 72 h values (Table 2.3). Prediction accuracy was quantified using R^2, MAE, and RMSE metrics; the accuracy results are explained in the following table (Table 3.2).

Table 3.2: Prediction accuracy metrics of 72-hour cell viability determined with LR models trained on 24-hour and 48-hour data with R^2, MAE, RMSE, and 95% CI values

Group	R^2	MAE	RMSE	95% CI (±)
A172 2D	0.79	8,80	11,35	±2.66
A172 3D	−0.13	15.56	17,62	±4.12
U87 2D	0.77	10,33	13,04	±3.05
U87 3D	**0.85**	**5,18**	**6,66**	**±1.56**

In summary, the models demonstrated accurate predictive performance, with the highest accuracy observed in 3D cultures, suggesting that the temporal stability of drug response is improved in 3D systems.

CHAPTER 4: DISCUSSION

The integration of biomimetic 3D technologies and AI in analyzing the TME of GB holds substantial potential in cancer research. Calculations based on 3D volume are generally more accurate than those based on 2D projections, as they display the entire shape and size of cellular structures. AI-based approaches and biomimetic 3D models have already made substantial contributions to the understanding of GB biology. When used in combination, these approaches offer even higher potential for precise and comprehensive investigations of tumor behavior, treatment responses, and therapeutic development. AI is particularly well suited to this domain due to its capacity to analyze large and complex datasets, identify patterns, and generate predictive models that exceed human analytical capabilities. AI-based algorithms can effectively analyze large volumes of data generated from 3D culture systems in GB research. This method enables valuable insights into cellular behavior, interactions within the TME, and responses to various therapeutic interventions. These capabilities are particularly critical for GB, a tumor characterized by pronounced heterogeneity and therapeutic resistance. The application of AI in medicine has demonstrated substantial value across multiple clinical domains. Researchers can use AI to uncover new biomarkers and possible therapeutic targets. They can also use personalized biomimetic 3D tumor models to predict how each patient will respond to a specific treatment. These models give a full picture of each patient's unique condition.

Biomimetic 3D models have demonstrated higher accuracy in recapitulating the complex architecture of GB compared with conventional 2D cultures. 3D technologies provide a physiologically relevant platform for assessing crosstalk among cancer cells, immune components, stromal cells, and the ECM within the GB TME, thereby yielding more biologically accurate results. Techniques such as 3D bioprinting, spheroid cultures, and organoid models have provided valuable insights into tumor invasion, angiogenesis, and immune-evasion mechanisms. By combining these models with AI,

scientists can conduct efficient analysis, accelerating the discovery of effective therapeutic compounds and optimizing treatment regimens.

In the drug response prediction part of this research, the utilization of merely two time points, 24 hours, and 48 hours, to predict future reactions of A127 and U87 cells, in both 2D and biomimetic 3D configurations, unavoidably presupposes linear regression of drug-induced cytotoxicity. The prediction results demonstrated adequate predictive accuracy at low to moderate TMZ doses; however, significant discrepancies were noted at elevated concentrations. Such variations indicate saturation effects, delayed cytotoxic responses, or inherently non-linear cell death dynamics that a simple linear model cannot adequately represent. The appearance of these discrepancies underscores biologically significant limitations of linear extrapolation rather than technical flaws, emphasizing the complexity of TMZ response kinetics. This reinforces the overall interpretation of the results.

While AI and 3D technologies exhibit potential, we still need to overcome several obstacles. The intricate nature of the datasets produced by biomimetic models poses a significant challenge. These systems frequently generate complex data encompassing spatial, temporal, molecular, and genetic information, posing challenges for their integration into comprehensive models. AI algorithms may require further refinement to effectively handle and interpret complex datasets. In addition, the accuracy of AI predictions is limited by the inherent variability of GB tumors, including the variations in genetic and phenotypic profiles observed among patients. Hence, enhancing the training of AI methods to consider the diversity of tumoroid GB research will be crucial in attaining improved patient-specific results.

One additional hurdle involves the *in vivo* validation of AI predictions. While AI may produce predictions utilizing *in vitro* biomimetic 3D models, it is crucial to validate these results in real clinical environments. Translating research findings from the laboratory to clinical application is a laborious and resource-demanding process. Prior to implementing AI-driven predictions in practical treatments, they must undergo rigorous testing to verify their safety and efficacy. Furthermore, it is imperative to meticulously assess the ethical and regulatory implications of employing AI-based

decision-making in clinical settings, particularly in the context of very aggressive tumors like GB.

In the future, AI in GB research is expected to focus on developing advanced models that integrate molecular and spatial data with patient-specific clinical data, such as imaging and genomic information. AI systems have the potential to continuously learn from changing datasets, enhancing their predictive capabilities and adaptability to the intricacies of GB biology.

The research utilizing AI and GB illustrated how the integration of 3D *in vitro* technology with AI datasets can transform our understanding of the multiplex 3D TME. This technology enabled us to capture multiple images simultaneously and under the same conditions, which were like those found in a tumor, while mimicking that environment on a microscopic scale. This type of combination of cell culture and AI could change how personalized cancer treatment and clinical decision-making evolve. However, it is crucial to continually improve ML models and perform cross-validation to ensure the predictions are accurate. In this series of experiments integrated with AI methodology, we can quantify specific parameters applicable for drug testing, precision medicine, and tumor segmentation.

This research demonstrated the ability of ML models to predict GB cell dynamics over time by analyzing confocal imaging data from 3D cell culture environments. By combining biomimetic 3D models and time series ML methods, specifically CNNs for extracting image features and RNN-based architectures, we were able to make predictions about essential biological parameters like migration capacity, proliferation rate, invasion rate, and morphology of cells. The GRU model exhibited superior predictive accuracy among those assessed, particularly when hyperparameters, such as the number of densities, were optimized, confirming the hypothesis that cellular behavior exhibits identifiable temporal patterns that are amenable to capture by sequential models.

The precise forecasting of GB cell behavior suggests that AI can uncover hidden dynamics within the TME, a complex system that influences tumor progression, therapeutic resistance, and metastasis. The use of biomimetic 3D cell cultures makes *in*

vitro conditions more similar to those *in vivo*. The integration of ML approaches enables the real-time observation and analysis of cellular processes. These technologies reduce reliance on animal models and accelerate preclinical drug screening, thereby enhancing the assessment of tumor responses to therapies and environmental influences, and ultimately improving precision medicine.

Similar studies in the literature enhance the interpretability of saliency-based DL models for glioma survival prediction by integrating domain knowledge derived from saliency maps. For research reference, a 3D CNN was utilized to categorize 147 patients' preoperative multimodal MRI scans from the BraTS 2020 dataset into groups based on short-term, medium-term, and long-term survival outcomes. By extending Grad-CAM to 3D and aligning saliency maps with an anatomical atlas, the method found clinically significant spatial patterns that linked tumor sites and volumes with overall survival outcomes. In the end, researchers have found that a clear link existed between short, medium, and extended survival times and tumors according to brain areas, such as the pallidum, fusiform, and transverse temporal gyrus. This anatomically informed visualization framework not only confirmed that larger tumor volumes correlate with a poorer prognosis but also provided clinicians with interpretable regional insights that facilitate the comprehension of model predictions, thus reconciling algorithmic reasoning with clinical decision-making (Osadebey, Liu, Fuster-Garcia, & Emblem, 2023).

This research expanded the generalizability and interpretability of artificial intelligence models by incorporating a biomimetic 3D cell culture GB TME confocal image dataset into deep learning methodologies for parameter predictions. Our study involved the culture and imaging of tumoroids produced from GB and astrocytes using confocal microscopy to evaluate migration and invasion behaviors under distinct microenvironmental circumstances. The application of ML algorithms to these data enabled a rapid and objective elucidation of disease mechanisms and the exploration of therapeutic potential through patient-derived drug screening. In this context, the RNN-LSTM hybrid model was employed, resulting in diminished error values and enhanced alignment with experimental data, thereby demonstrating the effectiveness of recurrent architectures in predicting normalized migration capacity. Nonetheless, the limited dataset and the need for further optimizations highlighted the importance of expanding

data resources and improving *in silico* frameworks to enhance predictive robustness, generalizability, and translational application.

Even with these improvements, several limitations remain. Despite the high predictive accuracy, we could achieve further improvements by using larger datasets and more diverse cell lines. Furthermore, incorporating multi-omics data, such as genomic or proteomic profiles, with imaging data could substantially enhance model robustness and biological interpretability.

AI has significant potential to enhance the clinical care of GB patients and precision medicine; however, several issues remain to be addressed, particularly in terms of execution. Cultivating cells from patients may prove more arduous due to their heterogeneous nature, complicating the examination of their interactions and behaviors.

The research demonstrated an innovative approach to enhance the evaluation of GB treatments and the attributes of the surrounding environment through the integration of 3D multiplex TME with a machine learning methodology. This combined approach may be effective for various types of cancer. This study also demonstrates the increasing importance of AI in 3D multiplex TME brain tumor imaging, which can enhance treatment planning and forecast a patient's future health. The ultimate objective of these biomimetic 3D TME studies combined with ML is precision medicine (Scimeca et al., 2021).

Another issue is that AI predictions need to be tested in patient-derived cells from living organisms. AI can make predictions based on 3D models cultured in TMEs; however, it is essential to confirm these results in real-life clinical settings. Translational research from the laboratory to clinical practice can be costly and may require additional financial support. Before AI-based predictions can be used in real-life treatments, they must be thoroughly tested to ensure they are safe and effective. Additionally, it is crucial to carefully consider the ethical and legal implications of using AI-based decision-making in clinical settings, particularly when treating patients with highly aggressive cancers, such as GB.

Future GB research will likely focus on integrative AI frameworks that combine molecular, geographical, and clinical datasets, thereby fostering adaptive models capable of encapsulating tumor heterogeneity and individual patient disease trajectories. These models will utilize advanced imaging techniques and genomic data to inform their development. AI systems can learn from datasets that evolve, which makes them more effective at making predictions and better adapted to the complexities of GB biology.

Researchers can utilize AI to identify new biomarkers and potential treatment targets, as well as to predict how each patient will respond to specific therapies. Personalized 3D tumor models facilitate comprehending each patient's distinct circumstances. Scientists can rapidly conduct a quick analysis by combining these models with AI, which may accelerate the discovery of beneficial therapeutic compounds and enhance treatment plans.

The integration of AI and 3D technologies in the analysis of the TME of GB presents significant potential in cancer research. RNN, LSTM, and GRU are effective ML models for analyzing time series data to understand how cells behave. The integration of AI with advanced 3D modeling techniques presents substantial potential for improving comprehension of the GB microenvironment and accelerating the formulation of tailored treatment strategies. As AI continues to evolve and integrate with biomimetic 3D models, there is a promising potential for more accurate and individualized approaches in GB treatment outcomes. To achieve more efficient outcomes, supplementary data may be utilized, and the dataset employed in the study can be enhanced, facilitating improved performance of deep learning models. 3D multiplexed environments are advantageous as they allow researchers to collect large amounts of data simultaneously within the same environmental conditions at a specific time.

Hybrid ML modes are used in the field of cancer research. The ML approach is employed for lung cancer segmentation and classification, whereas ViT methodologies are inspiring advancements in medical AI and cancer imaging. Recently, numerous scientists have devised ViT-based AI methodologies for lung cancer detection and

prognosis. DL algorithms are employed to detect early anomalies for cancer diagnosis (Murugesan et al., 2021).

The development of transformer-based models for EEG analysis has accelerated significantly. Researchers use transformers for robust architectures and applications in motor imagery, seizure detection, and emotion classification in epilepsy (Vafaei & Hosseini, 2025).

Engineering applications are diverse in terms of AI and hybrid ML models. A comparative analysis of the ML transformer model versus RNN and CNN-based models exists in the literature (Zhizheng Zhang et al., 2025). Transformers, as ML models, as well as CNN, DNN, and LSTM algorithms, are considered performant architectures for predictions (Barua, Haider, Austin, & Shawon, 2024).

Among the several DL methodologies, CNNs, RNNs, and transformer ML models have demonstrated significant efficacy in time-series data analysis and anomaly detection objectives. Scientists strive to select the most suitable and accurate model for various applications across various domains (Omoseebi et al., 2025).

Human Action Recognition (HAR) is a discipline that studies human behaviors across various fields. Convolutional Neural Networks (CNNs) have been utilized to improve performance, while ViTs are selected as an alternative (Alomar, Aysel, & Cai, 2024). Another study investigates three notable deep learning architectures, CNN, RNN, and ViT, for emotion recognition and compares the prediction results according to their accuracy (Yushchenko, Smelyakov, & Chupryna, 2025).

In this study, we observed that while transformer models are recognized for their high performance in various ML tasks, for our medium-sized data set input, their accuracy was lower compared to other deep learning models that we had utilized before, such as CNN, RNN, LSTM, and GRU. Specifically, for predicting migration capacity at day 7 using data from days 3 and 5, the RNN, its subtypes, LSTM, and GRU models achieved accuracies of `0.80`, `0.81`, and `0.83`, respectively, in our previous studies with the same input data for the ML model. In contrast, the transformer model did not perform

as well in this context when we fed the same input data to it. This mismatch might be attributed to the characteristics of our dataset.

CNN and RNN architectures are generally more performant, with higher accuracy, for relatively minor datasets and image-based inputs. At the same time, LSTMs and GRUs are particularly well-suited for modeling relatively short time series data. Transformers, on the other hand, typically require much larger datasets and are performant on very large databases to efficiently utilize their capacity for capturing complex relationships. As a result, the medium size of our dataset likely limited the performance of the transformer ML model, highlighting the importance of selecting ML architectures that align with the scale and nature of the available input data.

4.1 Complexity of data integration

The complexity and heterogeneity of the data generated constitute an important obstacle to applying AI to biomimetic models of the GB TME. 3D models often generate datasets with multiple dimensions, including spatial, temporal, and molecular data. These datasets necessitate the use of advanced algorithms for processing and interpretation. AI systems currently have limitations in fully capturing the complex and multifaceted interactions within the GB TME. Additionally, variations in results from various biomimetic models can create inconsistencies, complicating the establishment of standardized AI models (Thakuri, Liu, Luker, & Tavana, 2018).

AI and 3D technologies have considerable potential, but numerous issues still need to be addressed. Biomimetic models generate complex datasets, posing a significant challenge. These systems frequently provide intricate data encompassing geographical, temporal, molecular, and genetic information, complicating their integration into comprehensive models. AI algorithms may require further studies to handle and interpret complex datasets effectively. The accuracy of AI predictions is also limited by the fact that GB tumors are naturally variable, as evidenced by the discrepancies in genetic and phenotypic profiles observed in separate patients. Thus, augmenting the training of AI models to account for this diversity may be essential for achieving enhanced patient-specific outcomes.

Similar image precession studies apply segmentation of gliomas from MRI scans, which is essential for accurate diagnosis and treatment; however, the complexity of these tumors presents major barriers to advanced techniques. In the literature, we observe several studies that examine more DL models for glioma segmentation, emphasizing the integration of research with clinical application. Overall, three types of models exist: CNN-based, pure transformer, and hybrid. Each type is tested for its ability to segment, speed, and hardware requirements. A performance comparison of the top models is conducted, along with an analysis of their suitability. The researchers aim to highlight trends, constraints, and prospective research avenues that aim to augment technical performance and clinical applicability, thereby achieving accurate outcomes in glioma diagnosis (Diana-Albelda, García-Martín, & Bescos, 2025).

DNNs have significantly improved ML methods, particularly in processing images and natural language. CNN and RNN are two relevant architectures that can solve complex sequential problems with accuracy. Studies are comparing various RNN architectures, with a focus on LSTM and GRU models. Scientists apply particular models to estimate the force of pouring: a multi-layer LSTM, a multi-layer GRU, a single-layer LSTM, a single-layer GRU, and so forth (Nosouhian, Nosouhian, & Khoshouei, 2021).

Besides FL, the integration of spatial transcriptomics with time-series imaging data through multimodal AI models, especially transformer-based architectures, provides an effective method for simulating tumor progression and microenvironmental dynamics. By correlating regionally resolved gene expression patterns with longitudinal imaging techniques, such as MRI and live-cell microscopy, these methodologies encapsulate both molecular heterogeneity and temporal evolution (Al-Zoghby, Ebada, Saleh, Abdelhay, & Awad, 2025).

Multimodal transformers, engineered to simultaneously analyze graph-structured spatial omics and sequential imaging modalities, facilitate deep learning systems in comprehending cross-modality interactions, thereby enhancing phenotypic prediction, therapeutic response forecasting, and biomarker identification. This paradigm facilitates applications including forecasting the spatial arrangement of cellular responses to therapy and simulating the evolution of gene expression patterns over time, thereby advancing comprehensive, temporally informed precision oncology. The DL integration

with multimodal data improves cancer diagnosis and treatment, yielding more precise and individualized therapies (Rajdeo, Aronow, & Surya Prasath, 2024).

4.2 Computational resource requirements

The analysis of 3D tumor models using AI is highly demanding in terms of software and hardware resources, as well as significant computing power and sophisticated infrastructure. The handling, storage, and evaluation of the massive information produced by high-resolution imaging, omics data, and dynamic TME monitoring over time are crucial. Several research institutions require additional computational resources, which hinders the widespread adoption of AI in GB research (Ayensa-Jiménez, Doweidar, Sanz-Herrera, & Doblare, 2022).

Investment in research and development is essential for enhancing and constructing large-scale scientific facilities, which necessitate sophisticated algorithms and *in silico* methodologies. AI applications are facilitating transformative solutions for substantial data concerns; nevertheless, single-GPU solutions have become inadequate. AI integrates with high-performance computing (HPC) to reduce time-to-insight, enable domain-specific AI architectures, and optimize strategies for data-driven discovery. Recent advancements seek to expedite the utilization of HPC platforms for AI algorithms in both academia and industry (Huerta et al., 2020).

Some scientists examine the integration of HPC with cloud computing to enhance AI model training methodologies. They demonstrate the benefits of cloud-based HPC resources, encompassing scalability, flexibility, and cost-efficiency. The research additionally investigates diverse HPC architectures and frameworks for distributed training. Challenges include bottlenecks in data transfer, resource allocation, and energy usage (H. Sharma, 2019).

Creating and training AI models for GB research and profound learning systems for segmentation, molecular subtyping, and survival prediction requires significant computational resources, including advanced GPUs (e.g., NVIDIA Tesla K80, V100, A100), TPUs, and access to high-performance computing (HPC) clusters for extensive

parallel processing and model optimization. Various research teams employ cloud computing systems (e.g., Azure, AWS, and Google Cloud) to access resources with flexibility, such as Azure NV-series VMs equipped with multiple GPUs (Tesla M60 or V100/A100), 24–32 vCPUs, and 200–300 GB of RAM, which have proven effective for training CNNs and models based on ResNet or U-Net using GB MRI and genomic data (Goddla, 2022).

On a larger scale, FL research on GB, which included over 6.000 patients from 71 institutions and more than 25.000 MRI scans, used distributed training across global high-performance computing and cloud resources. This led to significant improvements in tumor boundary delineation without the need for centralized data sharing (Pati et al., 2022). The integration of AI with HPC infrastructures, encompassing hybrid cloud and on-premise clusters, is essential for facilitating big-data-driven discoveries in biomedical fields, including glioblastoma research (Huerta et al., 2020).

4.3 Data quality and validation of ML models

The effectiveness of AI algorithms is mainly dependent on the quality and quantity of data used for training the model. High-quality annotated datasets derived from 3D models are scarce in the context of GB. The restricted availability of data, particularly the absence of varied, well-annotated datasets encompassing underrepresented patient demographics, may lead to AI models being trained on incomplete or biased information (Tripathi et al., 2023), (Y. Yang et al.). The absence of inclusion across factors such as race, ethnicity, gender, age, and socioeconomic status may lead to systemic bias in models, yielding inaccurate or non-generalizable predictions (Marko, Neagu, & Anand, 2025).

Moreover, while AI may produce insights and predictions in controlled laboratory conditions, implementing these results in real-world and clinical contexts presents a significant hurdle. It is crucial to validate AI-derived findings in real clinical environments, particularly to guarantee that models are trained on datasets that encompass diverse, underrepresented patient populations and maintain equitable and reliable performance in practice (Zou & Schiebinger, 2021), (Melguizo-Gavilanes,

Bruner, Guha-Thakurta, Hess, & Puduvalli, 2015). A robust validation method involves clinical trials utilizing AI-assisted decision-making, such as randomized controlled trials where AI technologies aid in treatment selection, eligibility screening, or prognosis evaluation (Kang, Chowdhry, Pugh, & Park, 2023). These trials assist in confirming safety, efficacy, and generalizability within actual patient populations (J. Li et al., 2021).

4.4 Lack of standardization in biomimetic 3D models

The absence of standards in biomimetic 3D models is a barrier to generating universally accepted GB TME models. These models can vary in terms of scaffold materials, culture conditions, and cell sources. The most recent 3D *in vitro* models aim to bridge the gap between *in vitro* and *in vivo* environments. A trade-off between complexity and the standardization of the preparatory process persists (Piantino et al., 2021). Developing AI algorithms that can effectively apply to various systems is challenging due to the wide range of model types. For AI applications to achieve optimal effectiveness, it is crucial to establish a higher level of standardization and reproducibility in the development and characterization of 3D models. Without this essential factor, there is a high likelihood of substantial outcome variations across various research groups, thereby restricting the ability to compare AI-based analyses.

4.5 Biological representation limitations

Despite the progress made in 3D models, they still need to fully replicate the intricate *in vivo* nature of the GB TME. *In vitro* experiments often overlook or oversimplify relevant factors, such as immune cell interactions, systemic signaling, and vascular dynamics (Min Tang et al., 2024). AI predictions based on these simplified systems may not be biologically accurate enough to inform clinical decisions. Enhancing the biological relevance of AI-based research requires further refinements in the sophistication of biomimetic models. This method entails integrating more authentic tumor-immune interactions and vascular elements. An insightful and cohesive

understanding of the various cellular and molecular elements within the GB TME and their crosstalk is crucial for effective therapeutics (Bikfalvi et al., 2023).

4.6 Challenges in translating AI to clinical practice

Translating AI applications in 3D GB models from research settings to clinical practice presents serious challenges. Scientists emphasize the considerable potential of AI-driven ML and DL techniques in non-invasive medical assessment methods. The therapeutic importance of these AI-based advancements lies in their capacity to enhance GB patient care by providing accurate and timely information for treatment decisions. Implementing these technical developments in clinical practice presents challenges, including the need for extensive multi-institutional cohorts and the integration of various data formats. Regulatory and ethical considerations, as well as the requirement for thorough clinical validation, hinder the transition from laboratory findings to patient care (Samartha et al., 2024). Additionally, clinicians may require extensive training to effectively utilize AI-based tools in a clinical setting, and healthcare systems must adapt to integrate these technologies (Luo et al., 2023).

Challenges in image processing in neuroscience may stem from poor data quality, changes in experimental conditions, and the need for substantial computing power for advanced studies. These constraints may render the application of computational methodologies more challenging to validate *in vivo* and *in vitro* investigations, potentially reducing the reliability of the results. Still, computer-based image analysis is very useful, especially for getting new quantitative measures that describe how glioblastoma (GB) cells proliferate and migrate. These kinds of studies also help determine subcategories of diseases and the mechanisms behind them, which makes neuropathological classification more accurate. Similar studies exist for PD, as demonstrated in the following figure (Figure 4.1) (D'Sa et al., 2023). In the future, neuroscience-oriented image processing will involve the development of specialized software tools, enhanced predictive modeling for more accurate analytical results, and clinical decision-support systems that integrate data from multiple sources. These improvements may accelerate the implementation of precision medicine for GB patients.

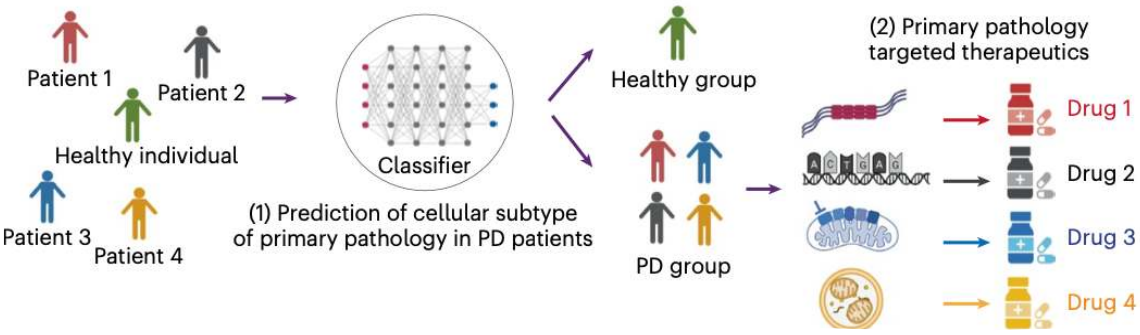

Figure 4.1: Usage of AI in the prediction of disease subtypes

4.7 Availability

Tumor heterogeneity, especially in GB, has traditionally been thought to lead to resistance to several therapies; however, recent evidence indicates this approach may be inaccurate. Clinical problems encompass neurologic and medical comorbidities in individuals with brain tumors, along with the potential toxicity associated with tumor-directed therapies (Barbaro, Fine, & Magge, 2021). Large, well-annotated datasets are crucial for AI research. Still, most GB data is siloed across institutions (Luo et al., 2023). Major data-sharing platforms, such as The Cancer Genome Atlas (TCGA-GB) via The Cancer Imaging Archive (TCIA), the International Cancer Genome Consortium (ICGC), and glioma-specific repositories like the Ivy Glioblastoma Atlas or RHUH-GB, have improved accessibility ("*Comprehensive genomic characterization defines human glioblastoma genes and core pathways*" 2008). Nonetheless, these repositories encounter constraints: TCGA-GB comprises a limited number of cases with diverse imaging protocols and frequently lacks longitudinal follow-up or uniform molecular annotation; multi-site datasets are constrained in postoperative or recurrence imaging and expert segmentations; clinical and demographic metadata may be incomplete or inconsistent, and local hospital cohorts are often underpowered or unrepresentative of larger populations (García-García, García-Galindo, Arrese, Sarabia, & Cepeda, 2022). Progress in artificial intelligence, particularly in machine learning and deep learning, has improved survival prediction for glioma patients by integrating multimodal data. Imaging-based models, non-imaging methodologies, and hybrid techniques are efficacious. Nevertheless, constraints include data heterogeneity, issues with

interpretability, and computational requirements. Proposed solutions include federated learning, lightweight AI models, and explainable AI frameworks (Awuah et al., 2025).

Furthermore, synthetic data production and data augmentation techniques, including Generative Adversarial Networks (GAN)- based MRI synthesis or mix-up-style patch mixing, may be suggested to mitigate the limitations of small datasets, especially when original patient data has restrictions or privacy issues that prevent its sharing (Shin et al., 2018). However, the untested robustness of AI models limits their application in GB (Dimitriadis, Trivizakis, Papanikolaou, Tsiknakis, & Marias, 2022).

To enhance AI resilience, it is essential to develop extensive, multicenter, harmonized datasets with uniform collection methodologies and comprehensive annotations, including various demographic information. However, the untested robustness of AI models continues to limit their application in GB. For AI models to achieve success, validation is essential, and multicenter methodologies are being investigated to confirm their efficacy. Data from GB patients indicate that innovative MRI techniques can assist in diagnosis, malignancy characterization, therapy planning, and assessment of therapeutic response (Nelson & Cha, 2003).

4.8 Interpretability

The absence of user-friendly software may hinder the clinical implementation of AI in GB, as primary users such as practitioners and radiologists are responsible for providing and communicating clinical diagnoses. Comprehensible AI is crucial for navigating the intricate field of healthcare decision-making (Walsh & Quail, 2023).

Ultimately, the potential of AI and 3D technology in enhancing GB research is exciting, but it is crucial to address the various challenges that may hinder their full impact. Enhancements in computational resources, data quality, model standardization, and validation of AI predictions in clinical settings are crucial for addressing these limitations.

4.9 Limitations

Several limitations prevent the broad application of the studies related to AI and GB in the articles that we have reviewed. Firstly, transcriptomic approaches rely on freshly frozen tissues, which are challenging to adapt to large-scale human research, as they typically use formalin-fixed paraffin-embedded (FFPE) tissues (Walsh & Quail, 2023). Secondly, due to their destructive nature, slow acquisition speeds, and limited multiplexing capability, several spatial proteomics techniques, such as Imaging Mass Cytometry (IMC), are less suitable for large-scale analysis. Cyclic immunohistochemistry (IHC) and immunofluorescence are two techniques that address tissue degradation and background signal accumulation over time (Luo et al., 2023). Getting large, well-annotated datasets, especially from randomized controlled trials, is a significant challenge when training AI models.

One of the major obstacles to integrating artificial intelligence with 3D technologies for GB research is the existence of data and modeling limitations. The establishment of a coculture system with GB cells to elucidate its invasive behavior represents another significant application of this paradigm. Notwithstanding these advantages, organoid models possess specific limitations, such as the lack of immunological responses and circulatory systems (Joseph, Blaavand, Daubon, Kruyt, & Thomsen, 2021). GB datasets frequently exhibit fragmentation and inconsistency, characterized by significant diversity in imaging techniques, genetic profiling methodologies, and annotation standards among institutions. This variation hinders the training of strong AI models and restricts their ability to generalize findings across diverse cohorts, ultimately undermining external validation and hindering clinical translation. Moreover, some AI frameworks, profound learning models, function as opaque black boxes, generating predictions without transparent reasoning or visibility into the underlying decision processes. Confronting these challenges requires standardized, multicenter data-sharing protocols and the development of explainable AI methodologies tailored to the biological complexities of GB. Overall, several significant barriers to clinical application remain, including inadequate data quantity and quality, the interpretability of algorithms, integration challenges, and ethical considerations (Alen Rončević, Nenad Koruga, Anamarija Soldo Koruga, & Robert Rončević, 2025).

Another major challenge in integrating AI with 3D technology for GB research is the quality and scalability of 3D tumor models. Numerous widely utilized platforms, such as Matrigel-based spheroids and organoids, may not accurately recreate the complete intricacy of the GB TME. These models frequently omit critical elements such as pericytes, microglia, neurons, immune cells, and functioning vasculature, providing only a limited representation of the brain's intrinsic ECM characteristics, such as stiffness, topography, and the hyaluronic acid-rich composition of the ECM. The self-organizing characteristics of organoids lead to diversity in structure and cellular composition, resulting in inadequate reproducibility.

Although organotypic brain slices are more physiologically relevant, their insufficient throughput renders them inappropriate for extensive drug screening. Composite hydrogels composed of Polyethylene Glycol (PEG), GelMA, and HAMA are utilized in 3D ECM models for tissue engineering and TME investigations. These hydrogels provide an equilibrium between structural integrity and bio-functionality, accommodating prolonged culture requirements and intricate cell-matrix interactions. This formulation is progressively used in complex 3D *in vitro* platforms to mimic *in vivo* TME accurately. In contrast, the hydrogels typically employed for 3D cultures have inherent limitations. Animal-derived matrices, such as Matrigel, exhibit batch-to-batch variability, whereas synthetic alternatives, like PEG or GelMA, often fail to adequately replicate the biomechanical and biochemical signals characteristic of the human brain. Such limitations severely hinder the translational relevance and scalability (Mier et al., 2025).

Additionally, the lack of user-friendly software and explainable AI models complicates clinical interpretation, and thorough testing of AI models in real-world scenarios is still required to ensure reliability (Cè et al., 2023). Concerns regarding potential overfitting in machine learning workflows are raised by the combination of abundant omics data and sparse drug response data, as well as the ongoing challenges in producing mature vascular structures in bio-printed models within experimental timelines (Min Tang et al., 2024).

Furthermore, the absence of a clear histological definition makes it more challenging to distinguish between pseudo-growth and genuine tumor progression. AI models designed for this purpose have not yet been extensively implemented in clinical practice. Additional research is required to incorporate these models into standard practice (Bikfalvi et al., 2023).

The promise of merging AI with 3D technologies in the study of GB TME is promising; however, certain limitations must be addressed to fully utilize their capabilities. By utilizing AI to assess the quality of organoid development, researchers can create more accurate *in vitro* models that closely resemble physiological conditions. AI is also essential for optimizing organoid bioprinting, guaranteeing precision and reproducibility. By combining AI with tumoroid technology, scientists will be able to overcome long-standing constraints, enhance experimental accuracy, and expedite clinical translation, all of which will help advance the field of health research (J. Yang, Fischer, & Ye, 2024).

Translational and ethical challenges provide significant obstacles to the clinical use of AI-enhanced 3D technologies in GB research. Despite progress in 3D tumor modeling and the implementation of patient-derived cell systems, these models frequently fail to reliably predict human therapeutic responses, thereby constraining their reliability for clinical decision-making (M. Tang, Rich, & Chen, 2021). Closing the divide between laboratory discoveries and practical applications remains a significant challenge, particularly when adapting these models for individualized treatment. Moreover, the advent of advanced technologies, such as bioprinting human-like brain tissues, raises intricate ethical, legal, and regulatory challenges. Unresolved concerns related to equal access to innovative medicines, standardization of production techniques, and oversight of bio-printed constructs. Regulatory frameworks have not yet matured sufficiently to comprehensively address these developments, hindering their clinical use, and generating significant concerns over safety, equity, and transparency in translational research (Wanigasekara, Cullen, Bourke, Tiwari, & Curtin, 2023).

In summary, using AI in the investigation of GB within 3D microenvironment models encounters constraints stemming from insufficient datasets, biases that promote racial and ethnic demographic data, and prolonged genetic drift. Deep learning models have a deficiency in interpretability, complicating the understanding of the reasoning behind their predictions. The absence of shared code for intricate AI structures, along with ethical issues about data utilization and patient confidentiality, complicates thorough research. The unpredictability of AI outputs remains a significant obstacle to clinical implementation in the GB research (T. Xie et al., 2024). Clinical translation for GB is impeded by biological and technological constraints, including the highly aggressive TME and resistance to conventional chemotherapy.

Conventional chemotherapeutic nanomedicine exhibits restricted clinical efficacy owing to inadequate cellular uptake, ineffective endosomal release, and tumor heterogeneity. Intricate delivery systems face challenges in medication dispersion, limited drug-loading capacity, and adverse effects. The persistent challenges in the GB microenvironment, along with anatomical, genetic, and cytological intricacies, hinder significant clinical advancement.

Overall, even with these limitations, the advantages of the 3D multiplex platforms remain beneficial in capturing a significant amount of data at once, as demonstrated in the following 3D confocal microscopy images (Figure 4.2, Figure 4.3), which illustrate GFP-tagged I-NHA human brain astrocytes and RFP-tagged human brain U87 GB cells in direct configuration obtained with multiplex ECM.

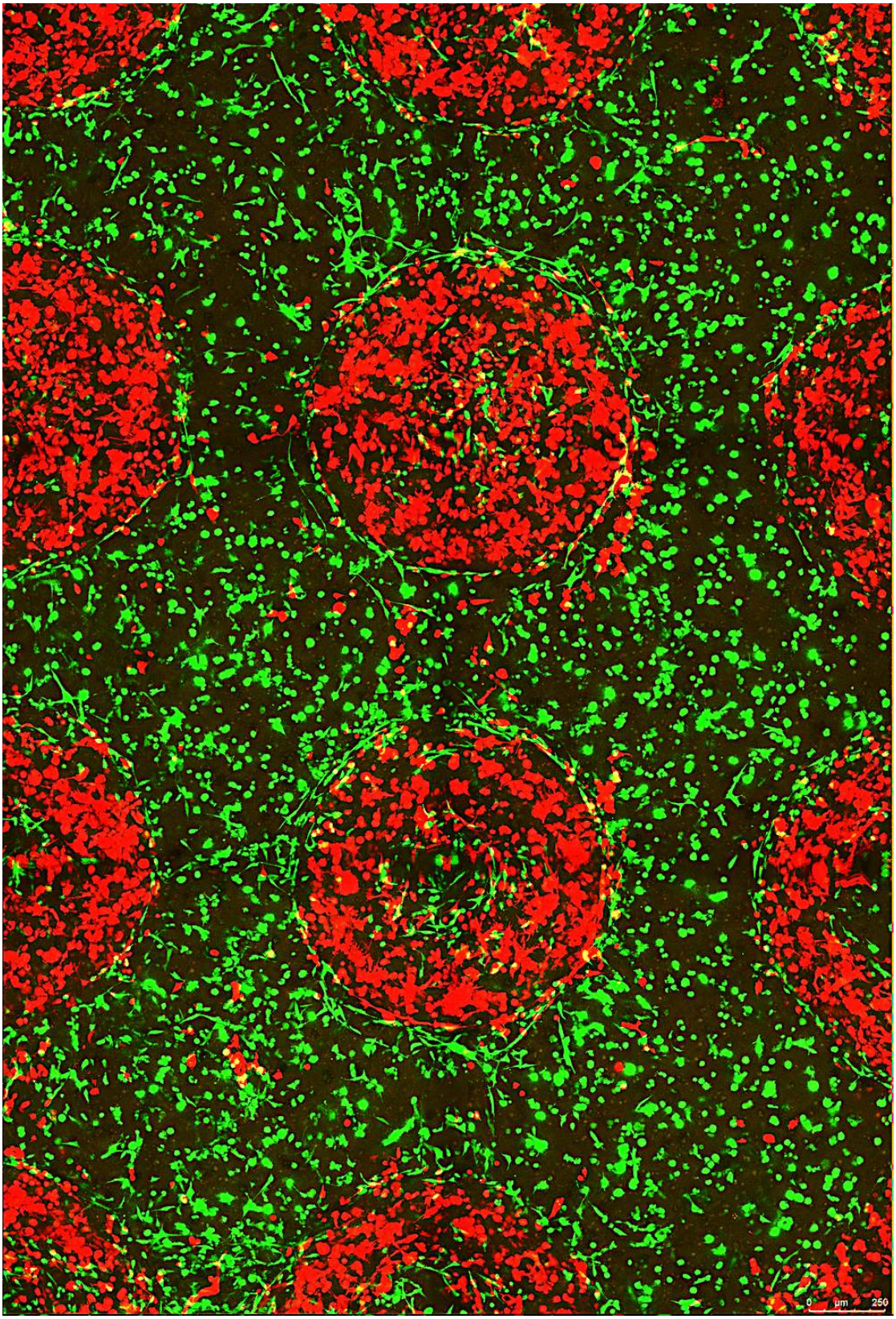

Figure 4.2: Tile scan of 3D confocal microscopy image of RFP-tagged U87 GB and GFP-tagged I-NHA astrocytes in stamping configuration obtained with multiplex ECM

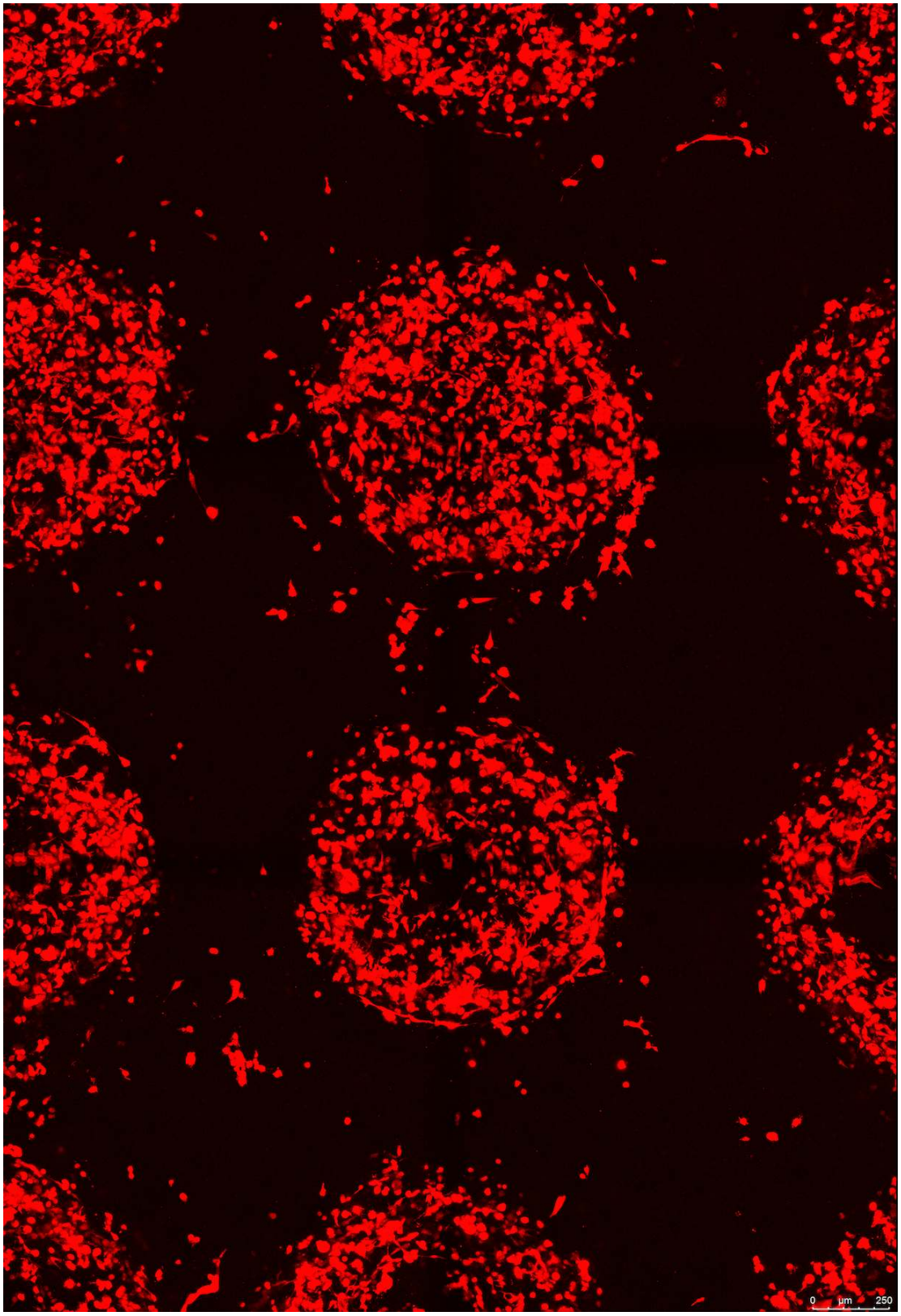

Figure 4.3: Tile scan of 3D confocal microscopy image of RFP-tagged human brain U87 GB cells in stamping co-culture configuration obtained with multiplex ECM

Transformer-based models and recurrent architectures such as LSTM and GRU differ substantially in their suitability for small, time-dependent biomedical datasets. Transformers typically require large datasets to learn meaningful patterns, and when trained on limited samples, they tend to underfit. This constraint is exacerbated by their absence of intrinsic sequential inductive bias, as transformers do not inherently encode temporal order; they rely on positional embeddings to infer progression from Day 3 to Day 5 to Day 7. When positional encoding is weak or insufficiently supported by the data, transformer models may not be the optimal fitting ML model to capture biological progression dynamics accurately; however, they perform well on large databases.

In contrast, LSTM and GRU architectures incorporate an intrinsic temporal flow through their recurrent hidden states, providing a strong inductive bias that aligns naturally with biological time-series input datasets. This architectural advantage makes them particularly well-suited for smaller image-derived datasets and short sequences, such as those used in predicting cell migration capacity and proliferation rate. Additionally, the computational characteristics of these models reinforce their generalizability. LSTMs and GRUs operate linearly across timesteps, rendering them both efficient and proportionate to raw data processing and prediction phases.

From a general perspective, biomedical datasets are often noisy, limited in size, and unevenly distributed in most cases. In such contexts, LSTM and GRU models tend to generalize more reliably because their more accurate, more constrained architectures mitigate overfitting and stabilize learning. On the other hand, Transformers typically excel only when large-scale, high-dimensional data with complex dependencies are available, such as in genomics or imaging datasets. Overall, in this study, these fundamental inequalities in data requirements, inductive biases, *in silico* structure, and generalization behavior explain why LSTM and GRU-based models outperform Transformers in our migration capacity and proliferation rate prediction task.

Ultimately, integration of AI and 3D cell culture technologies demonstrates significant potential in furthering our comprehension of the GB TME. Managing complex datasets and translating AI predictions into clinical practice poses challenges. Moreover, the prospective benefits of personalized medicine, drug discovery, and therapeutic optimization are substantial. The continuous advancement of AI combined with biomimetic 3D models holds significant promise for enhancing precision and personalized medicine in GB therapies.

CHAPTER 5: CONCLUSION

The application of AI in analyzing GB's biomimetic 3D microenvironments has increased recently, yielding promising outcomes despite the technology's limitations and challenges (Walsh & Quail, 2023). Further research is needed to realize AI's potential in GB management fully. Despite these gains, challenges persist in standardizing AI models to improve generalizability and interpretability (Luo et al., 2023).

AI-based models can help us better understand how brain tumor cell behave by analyzing time-series image data. AI also changes the way GB is managed by giving the possibility of diagnosis, treatment planning, and outcome prediction (Luo et al., 2023). AI-based models can enhance tumor segmentation, grading, and molecular characterization, thereby promoting personalized medicine strategies (Cè et al., 2023). Combining 3D multiplex TME with ML may enable the recreation of tumor characteristics and the prediction of therapeutic efficacy, which could lead to more treatment options, such as immunotherapy and anti-angiogenic therapies (Al-Rahbi et al., 2024). Moreover, AI applications in 3D technologies enhance our comprehension of the TME, which is essential for formulating targeted therapies (Walsh & Quail, 2023). Despite these advancements, challenges persist in standardizing AI models to enhance generalizability and interpretability, necessitating further research to harness AI's potential in GB management fully (Luo et al., 2023).

To mimic the complex nature of the GB TME, researchers have employed several 3D model systems, including spheroids, organoids, and bioprinted scaffolds. ML algorithms aim to understand how tumors work, predict how cells will behave, create personalized medicine, improve patient outcomes, and raise survival rates. The future advancement and widespread application of these technologies will advantage both physicians and patients, fostering precision medicine (Luger et al., 2022).

Our research demonstrates that the predictive accuracy of the ML models employed to assess migration capacity on day 7, utilizing data from days 3 and 5, was highly effective. Table 2 shows the full results. The consistently high values, all above 0.80, indicate that the training process was successful and that the models can identify considerable temporal patterns in the dataset. When the density number parameter in Python was changed from 10 to 1, the accuracy went up to 0.85. This observation illustrates the importance of tuning hyperparameters to enhance model performance. This finding indicates that a systematic examination of parameter configurations could significantly improve predicted results. Future research that expands this methodology to include additional hyperparameters, varied neural network architectures, and regularization methods may enhance prediction accuracy. These findings demonstrate the importance of optimization and methodological refinement in the development of predictive models for cell migration behaviors.

Early predictions were not always accurate enough, but over time, they became more accurate, showing how effective improved DL methods can be. Adding more datasets could make models even more reliable and accurate at making predictions. The results align with the extensive impact of AI in healthcare, where hybrid ML algorithms integrated with 3D cell culture platforms are revolutionizing diagnostics, refining personalized treatments, and improving patient care outcomes.

Advanced image processing and analysis with AI on biomimetic 3D cell culture platforms may significantly enhance the elucidation of cell interactions and behaviors, disease diagnosis, and patient treatment strategies, as well as improve our understanding of disease mechanisms.

In this study, we demonstrated the utility of AI algorithms in analyzing the biomimetic 3D TME of GB image datasets with a time series nature. AI-based methods enable the planning of personalized treatments and the prediction of how patients will respond to them, which is the primary goal of personalized medicine and treatment approaches. Combining AI with 3D technologies in TME analysis on multiplex cell culture platforms holds substantial promise for GB research, as multiplex platforms enable scientists to collect a large amount of data simultaneously. Biological replicates on

multiplex platforms provide enriched datasets, which ML algorithms can use as input to predict accurate results.

Various ML models, including CNNs, RNNs, and their subtypes, LSTM, and GRUs, are effective in predicting cell behaviors and interpreting time-series data. As an ML algorithm and a subtype of transformers, the ViT model may perform more accurately with larger training datasets, as transformers typically perform best with substantial amounts of high-quality data.

This investigation demonstrates that hybrid ML models, specifically RNN architectures and GRU models, can accurately forecast GB cell dynamics, including migration and proliferation behaviors, within a biomimetic 3D *in vitro* TME. Clinically, characterizing the temporal evolution of GB behavior in a 3D microenvironment provides valuable insight into invasion and recurrence patterns. The findings indicate that the incorporation of ML with multiplex 3D culture systems facilitates quantitative modeling of tumor activity, providing a robust framework for the analysis of cell-cell and cell-matrix interactions over time. While this research demonstrates encouraging predictive accuracy, it serves primarily as a proof of concept. This method yields more accurate, data-driven assessments of GB aggressiveness and therapy efficacy, in addition to technical performance, while allowing for the capture of multiple images simultaneously on multiplex platforms. Future investigations, including biomimetic patient-derived organoids, multi-omics datasets, and transformer-based designs, may enhance biological interpretability and translational applications, thereby facilitating AI-assisted, personalized therapy approaches in neuro-oncology and precision medicine.

In the TMZ drug response part of this research, linear regression ML models were utilized to generate temporal predictions of drug response results, comparing both CTRL and TMZ groups. They demonstrated significant accuracy in capturing dose-dependent trajectories over time. Overall, the results reveal clear and consistent discrepancies in cellular responses to TMZ treatment compared with the control condition across multiple time points. Control groups in both the A172 and U87 GB cell lines, cultured both in 2D and biomimetic 3D systems, demonstrated stable increases in viability, reflecting normal proliferative behavior. Machine learning predictions accurately extrapolate the 72-hour responses from wound closure from earlier temporal

measurements for both the TMZ and CTRL groups. An extended model of the MTT experiment predicted results for upcoming hours, 96, 120, and 144 hours, with the prediction results closely following the real data at 24, 48, and 72 hours, suggesting that early cellular responses may serve as reliable indicators of longer-term therapeutic outcome predictions in time-series data. Collectively, these findings highlight TMZ's robust inhibitory impact on GB cell viability and underscore the potential of early-timepoint ML models to anticipate long-term drug responses in GB research.

In summary, our results indicate that recurrent architectures such as LSTM and GRU outperform transformer-based models when predicting Day 7 migration outcomes from limited Day 3 and Day 5 biomedical data. Their superior performance stems from their data efficiency, inherent sequential structure, and suitability for short time-series tasks. Beyond model comparisons, the study highlights the growing role of AI in brain tumor research, where advanced *in silico* methods are poised to enhance treatment planning and support the long-term goal of personalized medicine. Integrating AI with emerging 3D technologies to analyze the GB TME highlights the transformative potential of these tools in cancer research. While CNNs, RNNs, LSTMs, and GRUs demonstrate strong capability for modeling cellular dynamics, ViTs may achieve comparable performance given larger and more diverse datasets. Ultimately, expanding the volume of available data will be crucial for enhancing predictive accuracy and fully leveraging the capabilities of modern machine learning approaches. The advantage of multiplex platforms is that they allow for the simultaneous capture of a large amount of data, including biological replicates.

In the future, obtaining more and higher-quality input data is crucial for achieving more accurate predictions. Additionally, we can address challenges like insufficient datasets and excessive memory consumption. Future research should focus on refining model parameters and feature selection to improve computational efficiency and predictive accuracy.

The integration of biomimetic 3D cell culture technologies with AI for analyzing the GB TME represents a promising advancement in cancer research. By strengthening the translational value of preclinical studies, these approaches may help ensure the biological relevance, safety, and efficacy of treatments prior to their clinical application.

This investigation underscores the growing importance of AI-based analysis in 3D multiplex brain tumor imaging, facilitating the prediction of cell interactions and behaviors on biomimetic 3D ECM models within specific hydrogels. Among machine learning approaches, gated RNN-based models demonstrated the highest accuracy in interpreting time-series cellular behaviors, while transformer-based architectures, such as ViT, may show potential with larger, high-quality datasets integrated with multi-omics data. Overall, continued improvements in data quality and model development are expected to accelerate progress toward personalized medicine and robust clinical decision-support systems.

In conclusion, this research illustrates the potential of integrating advanced ML techniques with image data obtained from biomimetic 3D GB models to predict cellular behavior within ECM multiplex platforms. The results indicate that RNN architectures, particularly GRU models, and their hybrid CNN variants, may effectively capture temporal dynamics in medium-sized image time series data, achieving robust predictive performance. While the findings demonstrate the potential for AI-based approaches to assist GB research and elucidate disease mechanisms and treatments, they also point to the importance of more extensive and diverse datasets, as well as standardized analytical frameworks. To transform these proof-of-concept results into reliable and widely usable instruments for preclinical applications and precision medicine, future endeavors need to concentrate on data augmentation, model interpretability, and cross-platform validation.

BIBLIOGRAPHY

Abdel Razek, A. A. K., Alksas, A., Shehata, M., AbdelKhalek, A., Abdel Baky, K., El-Baz, A., & Helmy, E. (2021). Clinical applications of artificial intelligence and radiomics in neuro-oncology imaging. *Insights Imaging, 12*(1), 152. doi:10.1186/s13244-021-01102-6

Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis, 22*(5), 717-727.

Akbari, H., Kazerooni, A. F., Ware, J. B., Mamourian, E., Anderson, H., Guiry, S., . . . Brem, S. (2021). Quantification of tumor microenvironment acidity in glioblastoma using principal component analysis of dynamic susceptibility contrast enhanced MR imaging. *Scientific reports, 11*(1), 15011.

Al-Rahbi, A., Al-Mahrouqi, O., & Al-Saadi, T. (2024). Uses of artificial intelligence in glioma: A systematic review. *Med Int (Lond), 4*(4), 40. doi:10.3892/mi.2024.164

Al-Zoghby, A. M., Ebada, A. I., Saleh, A. S., Abdelhay, M., & Awad, W. A. (2025). A Comprehensive Review of Multimodal Deep Learning for Enhanced Medical Diagnostics.

Ali, Z. (2024). A Comprehensive Overview and Comparative Analysis of CNN, RNN-LSTM and Transformer. *RNN-LSTM and Transformer (December 31, 2024)*.

Alieva, M., Barrera Román, M., de Blank, S., Petcu, D., Zeeman, A. L., Dautzenberg, N. M. M., . . . Rios, A. C. (2024). BEHAV3D: a 3D live imaging platform for comprehensive analysis of engineered T cell behavior and tumor response. *Nat Protoc, 19*(7), 2052-2084. doi:10.1038/s41596-024-00972-6

Alnawafleh, T. M., Radzi, Y., Alshipli, M., Oglat, A. A., & Alflahat, A. (2024). A Comprehensive Review of the Recent Advancements in Imaging Segmentation and Registration Techniques for Glioblastoma and Focusing on the Utilization of Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) Scans. *Curr Med Imaging, 20*(1), e15734056309829. doi:10.2174/0115734056309829240909095801

Alomar, K., Aysel, H. I., & Cai, X. (2024). RNNs, CNNs and transformers in human action recognition: A survey and a hybrid model. *arXiv preprint arXiv:2407.06162*.

Alpaydin, E. (2021). *Machine learning*: MIT press.

Amri, Y., Ben Slama, A., Mbarki, Z., Selmi, R., & Trabelsi, H. (2025). Automatic glioma segmentation based on efficient U-net model using MRI images. *Intelligence-Based Medicine, 11*, 100216. doi:https://doi.org/10.1016/j.ibmed.2025.100216

Andreatta, F., Beccaceci, G., Fortuna, N., Celotti, M., De Felice, D., Lorenzoni, M., . . . Alaimo, A. (2020). The Organoid Era Permits the Development of New Applications to Study Glioblastoma. *Cancers (Basel), 12*(11). doi:10.3390/cancers12113303

Atat, O. E., Farzaneh, Z., Pourhamzeh, M., Taki, F., Abi-Habib, R., Vosough, M., & El-Sibai, M. (2022). 3D modeling in cancer studies. *Hum Cell, 35*(1), 23-36. doi:10.1007/s13577-021-00642-9

Awuah, W. A., Ben-Jaafar, A., Roy, S., Nkrumah-Boateng, P. A., Tan, J. K., Abdul-Rahman, T., & Atallah, O. (2025). Predicting survival in malignant glioma using artificial intelligence. *European Journal of Medical Research, 30*(1), 61. doi:10.1186/s40001-025-02339-3

Ayensa-Jiménez, J., Doweidar, M. H., Sanz-Herrera, J. A., & Doblare, M. (2022). Understanding glioblastoma invasion using physically-guided neural networks with internal variables. *PLOS Computational Biology, 18*(4), e1010019. doi:10.1371/journal.pcbi.1010019

Badai, J., Bu, Q., & Zhang, L. (2020). Review of Artificial Intelligence Applications and Algorithms for Brain Organoid Research. *Interdisciplinary Sciences: Computational Life Sciences, 12*(4), 383-394. doi:10.1007/s12539-020-00386-4

Bai, L., Wu, Y., Li, G., Zhang, W., Zhang, H., & Su, J. (2024). AI-enabled organoids: Construction, analysis, and application. *Bioactive Materials, 31*, 525-548. doi:https://doi.org/10.1016/j.bioactmat.2023.09.005

Barbaro, M., Fine, H. A., & Magge, R. S. (2021). Scientific and Clinical Challenges within Neuro-Oncology. *World Neurosurg, 151*, 402-410. doi:10.1016/j.wneu.2021.01.151

Barua, A., Haider, S. N., Austin, S., & Shawon, S. M. (2024, 20-22 Dec. 2024). *Super Hybrid Deep Learning Framework: Transformer-LSTM-CNN-DNN for*

Accurate Load Forecasting. Paper presented at the 2024 27th International Conference on Computer and Information Technology (ICCIT).

Bhattacharya, T., Kumari, M., Kaur, K., Kaity, S., Arumugam, S., Ravichandiran, V., & Roy, S. (2024). Decellularized extracellular matrix-based bioengineered 3D breast cancer scaffolds for personalized therapy and drug screening. *J Mater Chem B, 12*(36), 8843-8867. doi:10.1039/d4tb00680a

Bikfalvi, A., da Costa, C. A., Avril, T., Barnier, J. V., Bauchet, L., Brisson, L., . . . Virolle, T. (2023). Challenges in glioblastoma research: focus on the tumor microenvironment. *Trends Cancer, 9*(1), 9-27. doi:10.1016/j.trecan.2022.09.005

Bittner, M. I., & Farajnia, S. (2022). AI in drug discovery: Applications, opportunities, and challenges. *Patterns (N Y), 3*(6), 100529. doi:10.1016/j.patter.2022.100529

Boehm, K. M., Aherne, E. A., Ellenson, L., Nikolovski, I., Alghamdi, M., Vázquez-García, I., . . . Shah, S. P. (2022). Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer. *Nat Cancer, 3*(6), 723-733. doi:10.1038/s43018-022-00388-9

Branco, F., Cunha, J., Mendes, M., Sousa, J. J., & Vitorino, C. (2025). 3D Bioprinting Models for Glioblastoma: From Scaffold Design to Therapeutic Application. *Advanced Materials, 37*(18), 2501994. doi:https://doi.org/10.1002/adma.202501994

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., & Jemal, A. (2024). Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin, 74*(3), 229-263. doi:10.3322/caac.21834

Breen, W. G., Aryal, M. P., Cao, Y., & Kim, M. M. (2024). Integrating multi-modal imaging in radiation treatments for glioblastoma. *Neuro Oncol, 26*(12 Suppl 2), S17-s25. doi:10.1093/neuonc/noad187

Burgstaller, G., Oehrle, B., Koch, I., Lindner, M., & Eickelberg, O. (2013). Multiplex Profiling of Cellular Invasion in 3D Cell Culture Models. *PLOS ONE, 8*(5), e63121. doi:10.1371/journal.pone.0063121

Cai, Y., Li, W., Zahid, T., Zheng, C., Zhang, Q., & Xu, K. (2023). Early prediction of remaining useful life for lithium-ion batteries based on CEEMDAN-transformer-DNN hybrid model. *Heliyon, 9*(7). doi:10.1016/j.heliyon.2023.e17754

Cè, M., Irmici, G., Foschini, C., Danesini, G. M., Falsitta, L. V., Serio, M. L., . . . Cellina, M. (2023). Artificial Intelligence in Brain Tumor Imaging: A Step toward Personalized Medicine. *Current Oncology, 30*(3), 2673-2701. doi:10.3390/curroncol30030203

Cellina, M., Cè, M., Irmici, G., Ascenti, V., Caloro, E., Bianchi, L., . . . Carrafiello, G. (2022). Artificial intelligence in emergency radiology: where are we going? *Diagnostics, 12*(12), 3223.

Charalel, R. A., Lin, H. S., & Singh, K. (2014). Glaucoma screening using relative afferent pupillary defect. *J Glaucoma, 23*(3), 169-173. doi:10.1097/IJG.0b013e31826a9742

Chatzianastasis, M., Vazirgiannis, M., & Zhang, Z. (2023). Explainable Multilayer Graph Neural Network for cancer gene prediction. *Bioinformatics, 39*(11). doi:10.1093/bioinformatics/btad643

Chen, L., Qin, D., Guo, X., Wang, Q., & Li, J. (2021). Putting Proteomics Into Immunotherapy for Glioblastoma. *Front Immunol, 12*, 593255. doi:10.3389/fimmu.2021.593255

Chen, P., Wang, P., & Gao, B. (2024). The application value of deep learning in the background of precision medicine in glioblastoma. *Sci Prog, 107*(1), 368504231223353. doi:10.1177/00368504231223353

Choi, J. W., Koo, D. L., Kim, D. H., Nam, H., Lee, J. H., Hong, S.-N., & Kim, B. (2024). A novel deep learning model for obstructive sleep apnea diagnosis: hybrid CNN-Transformer approach for radar-based detection of apnea-hypopnea events. *Sleep, 47*(12), zsae184. doi:10.1093/sleep/zsae184

Choi, R. Y., Coyner, A. S., Kalpathy-Cramer, J., Chiang, M. F., & Campbell, J. P. (2020). Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl Vis Sci Technol, 9*(2), 14. doi:10.1167/tvst.9.2.14

Choi, S. H., Kim, Y. H., Hebisch, M., Sliwinski, C., Lee, S., D'Avanzo, C., . . . Kim, D. Y. (2014). A three-dimensional human neural cell culture model of Alzheimer's disease. *Nature, 515*(7526), 274-278. doi:10.1038/nature13800

Chua, J., Nafziger, E., & Leung, D. (2019). Evidence-Based Practice: Temozolomide Beyond Glioblastoma. *Curr Oncol Rep, 21*(4), 30. doi:10.1007/s11912-019-0783-5

Clara-Trujillo, S., Marín-Payá, J. C., Cordón, L., Sempere, A., Gallego Ferrer, G., & Gómez Ribelles, J. L. (2019). Biomimetic microspheres for 3D mesenchymal

stem cell culture and characterization. *Colloids Surf B Biointerfaces, 177*, 68-76. doi:10.1016/j.colsurfb.2019.01.050

Comprehensive genomic characterization defines human glioblastoma genes and core pathways. (2008). *Nature, 455*(7216), 1061-1068. doi:10.1038/nature07385

Conte, L., Caruso, G., Philip, A. K., Cucci, F., De Nunzio, G., Cascio, D., & Caffo, M. (2025). Artificial Intelligence-Assisted Drug and Biomarker Discovery for Glioblastoma: A Scoping Review of the Literature. *Cancers, 17*(4). doi:10.3390/cancers17040571

D'Sa, K., Evans, J., Virdi, G., Vecchi, G., Adam, A., Bertolli, O., . . . Gandhi, S. (2023). Prediction of mechanistic subtypes of Parkinson's using patient-derived stem cell models. *Nature Machine Intelligence, 5*, 1-14. doi:10.1038/s42256-023-00702-9

Diana-Albelda, C., García-Martín, Á., & Bescos, J. (2025). A Review on Deep Learning Methods for Glioma Segmentation, Limitations, and Future Perspectives. *Journal of Imaging, 11*(8). doi:10.3390/jimaging11080269

Diao, W., Tong, X., Yang, C., Zhang, F., Bao, C., Chen, H., . . . Ou-Yang, Z. C. (2019). Behaviors of Glioblastoma Cells in in Vitro Microenvironments. *Sci Rep, 9*(1), 85. doi:10.1038/s41598-018-36347-7

Dimitriadis, A., Trivizakis, E., Papanikolaou, N., Tsiknakis, M., & Marias, K. (2022). Enhancing cancer differentiation with synthetic MRI examinations via generative models: a systematic review. *Insights into Imaging, 13*(1), 188. doi:10.1186/s13244-022-01315-3

Ding, J., Hu, X. H., & Gudivada, V. (2021). A Machine Learning Based Framework for Verification and Validation of Massive Scale Image Data. *IEEE Transactions on Big Data, 7*(2), 451-467. doi:10.1109/TBDATA.2017.2680460

Dulam, N., Gade, K. R., & Gosukonda, V. (2023). Generative AI for Data Augmentation in Machine Learning. *Journal of AI-Assisted Scientific Discovery, 3*(2), 665-688.

Dumbrăveanu, L., Cușnir, V., & Bobescu, D. (2021). A review of neovascular glaucoma. Etiopathogenesis and treatment. *Rom J Ophthalmol, 65*(4), 315-329. doi:10.22336/rjo.2021.66

Duval, K., Grover, H., Han, L. H., Mou, Y., Pegoraro, A. F., Fredberg, J., & Chen, Z. (2017). Modeling Physiological Events in 2D vs. 3D Cell Culture. *Physiology (Bethesda), 32*(4), 266-277. doi:10.1152/physiol.00036.2016

Elhanani, O., Ben-Uri, R., & Keren, L. (2023). Spatial profiling technologies illuminate the tumor microenvironment. *Cancer Cell, 41*(3), 404-420. doi:10.1016/j.ccell.2023.01.010

Enuganti, P. K., Sen Bhattacharya, B., Serrano Gotarredona, T., & Rhodes, O. (2025). Neuromorphic Computing and Applications: A Topical Review. *WIREs Data Mining and Knowledge Discovery, 15*(2), e70014. doi:https://doi.org/10.1002/widm.70014

Fan, J., Fang, L., Wu, J., Guo, Y., & Dai, Q. (2020). From Brain Science to Artificial Intelligence. *Engineering, 6*(3), 248-252. doi:https://doi.org/10.1016/j.eng.2019.11.012

Fan, Y., Sun, Y., Chang, W., Zhang, X., Tang, J., Zhang, L., & Liao, H. (2018). Bioluminescence imaging and two-photon microscopy guided laser ablation of GBM decreases tumor burden. *Theranostics, 8*(15), 4072-4085. doi:10.7150/thno.25357

Fangohr, H. (2004, 2004//). *A Comparison of C, MATLAB, and Python as Teaching Languages in Engineering.* Paper presented at the Computational Science - ICCS 2004, Berlin, Heidelberg.

Fathi Kazerooni, A., Kraya, A., Rathi, K. S., Kim, M. C., Vossough, A., Khalili, N., . . . Nabavizadeh, A. (2025). Multiparametric MRI along with machine learning predicts prognosis and treatment response in pediatric low-grade glioma. *Nature Communications, 16*(1), 340. doi:10.1038/s41467-024-55659-z

Feizi, A., & Gutierrez, J. M. (2022). Harnessing AI and Genomics to Accelerate Drug Discovery. In S. Ehsani, P. Glauner, P. Plugmann, & F. M. Thieringer (Eds.), *The Future Circle of Healthcare: AI, 3D Printing, Longevity, Ethics, and Uncertainty Mitigation* (pp. 89-106). Cham: Springer International Publishing.

Feng, Y., Zhao, H., Li, X., Zhang, X., & Li, H. (2017). A multi-scale 3D Otsu thresholding algorithm for medical image segmentation. *Digital Signal Processing, 60*, 186-199. doi:https://doi.org/10.1016/j.dsp.2016.08.003

Flores, J. E., Claborne, D. M., Weller, Z. D., Webb-Robertson, B. M., Waters, K. M., & Bramer, L. M. (2023). Missing data in multi-omics integration: Recent advances through artificial intelligence. *Front Artif Intell, 6*, 1098308. doi:10.3389/frai.2023.1098308

García-García, S., García-Galindo, M., Arrese, I., Sarabia, R., & Cepeda, S. (2022). Current Evidence, Limitations and Future Challenges of Survival Prediction for

Glioblastoma Based on Advanced Noninvasive Methods: A Narrative Review. *Medicina (Kaunas), 58*(12). doi:10.3390/medicina58121746

Gershman, S. (2024). What have we learned about artificial intelligence from studying the brain? *Biological Cybernetics, 118*, 1-5. doi:10.1007/s00422-024-00983-2

Ghadimi, D. J., Vahdani, A. M., Karimi, H., Ebrahimi, P., Fathi, M., Moodi, F., . . . Saligheh Rad, H. (2025). Deep Learning-Based Techniques in Glioma Brain Tumor Segmentation Using Multi-Parametric MRI: A Review on Clinical Applications and Future Outlooks. *J Magn Reson Imaging, 61*(3), 1094-1109. doi:10.1002/jmri.29543

Ghosh, S., & Herberg, S. (2024). ECM biomaterials for modeling of outflow cell biology in health and disease. *Biomater Biosyst, 13*, 100091. doi:10.1016/j.bbiosy.2024.100091

Goddla, V. (2022). *Artificial Intelligence Solution for Effective Treatment Planning for Glioblastoma Patients*.

Gönen, M., & Margolin, A. A. (2014). Localized data fusion for kernel k-means clustering with application to cancer biology. *Advances in neural information processing systems, 27*.

Goto, M., Suzuki, M., Mizukami, S., Abe, O., Aoki, S., Miyati, T., . . . Takeda, T. (2016). Repeatability of Brain Volume Measurements Made with the Atlas-based Method from T(1)-weighted Images Acquired Using a 0.4 Tesla Low Field MR Scanner. *Magn Reson Med Sci, 15*(4), 365-370. doi:10.2463/mrms.mp.2015-0107

Gulli, A., Kapoor, A., & Pal, S. (2019). *Deep learning with TensorFlow 2 and Keras: regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API*: Packt Publishing Ltd.

Guo, R., Chen, Y., Zhang, J., Zhou, Z., Feng, B., Du, X., . . . Cui, H. (2024). Neural Differentiation and spinal cord organoid generation from induced pluripotent stem cells (iPSCs) for ALS modelling and inflammatory screening. *Mol Neurobiol, 61*(7), 4732-4749. doi:10.1007/s12035-023-03836-4

Gurkan, U. A., Fan, Y., Xu, F., Erkmen, B., Urkac, E. S., Parlakgul, G., . . . Demirci, U. (2013). Simple precision creation of digitally specified, spatially heterogeneous, engineered tissue architectures. *Adv Mater, 25*(8), 1192-1198. doi:10.1002/adma.201203261

Gurusamy, N., Alsayari, A., Rajasingh, S., & Rajasingh, J. (2018). Adult stem cells for regenerative therapy. *Progress in molecular biology and translational science, 160*, 1-22.

Habanjar, O., Diab-Assaf, M., Caldefie-Chezet, F., & Delort, L. (2021). 3D Cell Culture Systems: Tumor Application, Advantages, and Disadvantages. *Int J Mol Sci, 22*(22). doi:10.3390/ijms222212200

Hasan, S. S., Fury, M. S., Woo, J. J., Kunze, K. N., & Ramkumar, P. N. (2024). Ethical Application of Generative Artificial Intelligence in Medicine. *Arthroscopy*. doi:10.1016/j.arthro.2024.12.011

Hinton, G., & Jordan, M. I. (2024). Advancing Healthcare, E-Commerce, and Computational Analysis with AI-Applications in Diagnostics, Market Insights, and Efficiency. *AlgoVista: Journal of AI and Computer Science, 1*(3), 592661.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation, 9*(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735

Hofmann, S., Cohen-Harazi, R., Maizels, Y., & Koman, I. (2022). Patient-derived tumor spheroid cultures as a promising tool to assist personalized therapeutic decisions in breast cancer. *Transl Cancer Res, 11*(1), 134-147. doi:10.21037/tcr-21-1577

Hu, Z., Wang, Z., Jin, Y., & Hou, W. (2023). VGG-TSwinformer: Transformer-based deep learning model for early Alzheimer's disease prediction. *Computer Methods and Programs in Biomedicine, 229*, 107291. doi:https://doi.org/10.1016/j.cmpb.2022.107291

Huang, K., Xiao, C., Glass, L. M., Zitnik, M., & Sun, J. (2020). SkipGNN: predicting molecular interactions with skip-graph networks. *Sci Rep, 10*(1), 21092. doi:10.1038/s41598-020-77766-9

Huerta, E., Khan, A., Davis, E., Bushell, C., Gropp, W., Katz, D., . . . Saxton, A. (2020). Convergence of artificial intelligence and high performance computing on NSF-supported cyberinfrastructure. *Journal of big data, 7*. doi:10.1186/s40537-020-00361-2

Huppertz, H. J., Kröll-Seger, J., Klöppel, S., Ganz, R. E., & Kassubek, J. (2010). Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures. *Neuroimage, 49*(3), 2216-2224. doi:10.1016/j.neuroimage.2009.10.066

Ismayilzada, N. (2024). MSc Thesis, The development of the 3D micropatterned model of the glioma-astrocyte co-culture: investigating glioma migration and astrocyte reactivity.

Issa, N. T., Stathias, V., Schürer, S., & Dakshanamurthy, S. (2021). Machine and deep learning approaches for cancer drug repurposing. *Semin Cancer Biol, 68*, 132-142. doi:10.1016/j.semcancer.2019.12.011

Jensen, C., & Teng, Y. (2020). Is It Time to Start Transitioning From 2D to 3D Cell Culture? *Frontiers in Molecular Biosciences, Volume 7 - 2020*. doi:10.3389/fmolb.2020.00033

Jiang, L., Niu, G., Liu, Y., Yu, W., Wu, H., Xie, Z., . . . Ren, W. (2021). Establishment and Verification of Neural Network for Rapid and Accurate Cytological Examination of Four Types of Cerebrospinal Fluid Cells. *Front Med (Lausanne), 8*, 749146. doi:10.3389/fmed.2021.749146

Joseph, J. V., Blaavand, M. S., Daubon, T., Kruyt, F. A., & Thomsen, M. K. (2021). Three-dimensional culture models to study glioblastoma - current trends and future perspectives. *Curr Opin Pharmacol, 61*, 91-97. doi:10.1016/j.coph.2021.08.019

Jung, A. (2022). *Machine learning: the basics*: Springer Nature.

Kang, J., Chowdhry, A. K., Pugh, S. L., & Park, J. H. (2023). Integrating Artificial Intelligence and Machine Learning Into Cancer Clinical Trials. *Seminars in Radiation Oncology, 33*(4), 386-394. doi:https://doi.org/10.1016/j.semradonc.2023.06.004

Kc, K., Li, R., Cui, F., & Haake, A. R. (2022). Predicting Biomedical Interactions With Higher-Order Graph Convolutional Networks. *IEEE/ACM Trans Comput Biol Bioinform, 19*(2), 676-687. doi:10.1109/tcbb.2021.3059415

Kim, H. Y., Lee, B. I., Jeon, J. H., Kim, D. K., Kang, S. G., Shim, J. K., . . . Jang, H. (2019). Gossypol Suppresses Growth of Temozolomide-Resistant Glioblastoma Tumor Spheres. *Biomolecules, 9*(10). doi:10.3390/biom9100595

Kong, J., Lee, H., Kim, D., Han, S. K., Ha, D., Shin, K., & Kim, S. (2020). Network-based machine learning in colorectal and bladder organoid models predicts anti-cancer drug efficacy in patients. *Nat Commun, 11*(1), 5485. doi:10.1038/s41467-020-19313-8

Kostadinov, S. (2018). *Recurrent neural networks with Python quick start guide: sequential learning and language modeling with tensorflow*: Packt Publishing Ltd.

Kramer, N., Walzl, A., Unger, C., Rosner, M., Krupitza, G., Hengstschläger, M., & Dolznig, H. (2013). In vitro cell migration and invasion assays. *Mutat Res, 752*(1), 10-24. doi:10.1016/j.mrrev.2012.08.001

Krieger, T. G., Tirier, S. M., Park, J., Jechow, K., Eisemann, T., Peterziel, H., . . . Conrad, C. (2020). Modeling glioblastoma invasion using human brain organoids and single-cell transcriptomics. *Neuro Oncol, 22*(8), 1138-1149. doi:10.1093/neuonc/noaa091

Krishnan, N. A., Kodamana, H., & Bhattoo, R. (2024). Model Refinement. In *Machine Learning for Materials Discovery: Numerical Recipes and Practical Applications* (pp. 131-143): Springer.

Labatut, V., Pastor, J., & Ruff, S. (2002). *Le cerveau humain: un réseau causal fonctionnel?* Paper presented at the 1er Colloque des doctorants de l'École Doctorale Informatique et Télécommunications (EDIT).

Lan, Y.-L., Zou, S., Qin, B., & Zhu, X. (2023). Potential roles of transformers in brain tumor diagnosis and treatment. *Brain-X, 1*(2), e23. doi:https://doi.org/10.1002/brx2.23

Lancaster, M. A., & Knoblich, J. A. (2014). Organogenesis in a dish: modeling development and disease using organoid technologies. *science, 345*(6194), 1247125. doi:10.1126/science.1247125

Learning, S.-S. (2006). Semi-supervised learning. *CSZ2006. html, 5*, 2.

Lee, J. H., Daugharthy, E. R., Scheiman, J., Kalhor, R., Ferrante, T. C., Terry, R., . . . Aach, J. (2015). Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nature protocols, 10*(3), 442-458.

Lee, J. H., Daugharthy, E. R., Scheiman, J., Kalhor, R., Yang, J. L., Ferrante, T. C., . . . Amamoto, R. (2014). Highly multiplexed subcellular RNA sequencing in situ. *science, 343*(6177), 1360-1363.

Lee, S. H., Shim, K. Y., Kim, B., & Sung, J. H. (2017). Hydrogel-based three-dimensional cell culture for organ-on-a-chip applications. *Biotechnol Prog, 33*(3), 580-589. doi:10.1002/btpr.2457

Li, H. L., Han, P. H., Pan, D. Q., Chen, G., Lu, X. H., & Li, J. (2020). LncRNA XIST regulates cell proliferation, migration and invasion of glioblastoma via

regulating miR-448 and ROCK1. *J Biol Regul Homeost Agents, 34*(6), 2049-2058. doi:10.23812/20-558-l

Li, J., Wu, J., Zhao, Z., Zhang, Q., Shao, J., Wang, C., . . . Li, W. (2021). Artificial intelligence-assisted decision making for prognosis and drug efficacy prediction in lung cancer patients: a narrative review. *J Thorac Dis, 13*(12), 7021-7033. doi:10.21037/jtd-21-864

Liu, J., & Wang, B. (2024). Application of 3D-Slicer Software in the Treatment of Gliomas. *J Craniofac Surg*. doi:10.1097/scs.0000000000010786

Lord, E. M., Penney, D. P., Sutherland, R. M., & Cooper, R. A., Jr. (1979). Morphological and functional characteristics of cells infiltrating and destroying tumor multicellular spheroids in vivo. *Virchows Arch B Cell Pathol Incl Mol Pathol, 31*(2), 103-116. doi:10.1007/bf02889928

Louit, A., Galbraith, T., & Berthod, F. (2023). In Vitro 3D Modeling of Neurodegenerative Diseases. *Bioengineering (Basel), 10*(1). doi:10.3390/bioengineering10010093

Luger, A. L., König, S., Samp, P. F., Urban, H., Divé, I., Burger, M. C., . . . Ronellenfitsch, M. W. (2022). Molecular matched targeted therapies for primary brain tumors-a single center retrospective analysis. *J Neurooncol, 159*(2), 243-259. doi:10.1007/s11060-022-04049-w

Luo, J., Pan, M., Mo, K., Mao, Y., & Zou, D. (2023). Emerging role of artificial intelligence in diagnosis, classification and clinical management of glioma. *Seminars in Cancer Biology, 91*, 110-123. doi:https://doi.org/10.1016/j.semcancer.2023.03.006

Macpherson, T., Churchland, A., Sejnowski, T., DiCarlo, J., Kamitani, Y., Takahashi, H., & Hikida, T. (2021). Natural and Artificial Intelligence: A brief introduction to the interplay between AI and neuroscience research. *Neural Networks, 144*, 603-613.

Maddineni, S., Janapati, S., Kosana, V., & Teeparthi, K. (2022, 19-20 Nov. 2022). *A hybrid deep transformer model for epileptic seizure prediction.* Paper presented at the 2022 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC).

Madonna, M. C., Duer, J. E., McKinney, B. J., Sunassee, E. D., Crouch, B. T., Ilkayeva, O., . . . Ramanujam, N. (2022). In vivo metabolic imaging identifies lipid

vulnerability in a preclinical model of Her2+/Neu breast cancer residual disease and recurrence. *NPJ Breast Cancer, 8*(1), 111.

Mancino, R., Martucci, A., Cesareo, M., Giannini, C., Corasaniti, M. T., Bagetta, G., & Nucci, C. (2018). Glaucoma and Alzheimer Disease: One Age-Related Neurodegenerative Disease of the Brain. *Curr Neuropharmacol, 16*(7), 971-977. doi:10.2174/1570159x16666171206144045

Mansouri, S. S., Sivaram, A., Savoie, C. J., & Gani, R. (2025). Models, modeling and model-based systems in the era of computers, machine learning and AI. *Computers & Chemical Engineering, 194*, 108957. doi:https://doi.org/10.1016/j.compchemeng.2024.108957

Marko, J. G. O., Neagu, C. D., & Anand, P. B. (2025). Examining inclusivity: the use of AI and diverse populations in health and social care: a systematic review. *BMC Medical Informatics and Decision Making, 25*(1), 57. doi:10.1186/s12911-025-02884-1

Massih, B., Veh, A., Schenke, M., Mungwa, S., Seeger, B., Selvaraj, B. T., . . . Lüningschrör, P. (2023). A 3D cell culture system for bioengineering human neuromuscular junctions to model ALS. *Front Cell Dev Biol, 11*, 996952. doi:10.3389/fcell.2023.996952

Melguizo-Gavilanes, I., Bruner, J. M., Guha-Thakurta, N., Hess, K. R., & Puduvalli, V. K. (2015). Characterization of pseudoprogression in patients with glioblastoma: is histology the gold standard? *Journal of neuro-oncology, 123*, 141-150.

Mier, M. S. A., Türker, E., Faber, J., Friedrich, M., Lamberger, Z., Weigelt, J., . . . Villmann, C. (2025). 3D In Vitro Glioma-Neuron-Astrocyte Biomimetic Composites Recapitulate Key Molecular Mechanisms Linked to Glioblastoma Multiforme Pathophysiology. *Advanced Functional Materials, 35*(22), 2419211. doi:https://doi.org/10.1002/adfm.202419211

Mnich, K., Lhomond, S., Wallace, E., Le Reste, P.-J., Pandit, A., Chevet, E., . . . Gorman, A. M. (2025). Improving glioblastoma treatment with imaging, radiotherapy, drug delivery, and therapeutic systems. *Device, 3*(2). doi:10.1016/j.device.2024.100685

Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python: a guide for data scientists*: " O'Reilly Media, Inc.".

Murugesan, M., Kaliannan, K., Balraj, S., Singaram, K., Kaliannan, T., & Albert, J. R. (2021). A Hybrid deep learning model for effective segmentation and

classification of lung nodules from CT images. *Journal of Intelligent & Fuzzy Systems, 42*(3), 2667-2679. doi:10.3233/JIFS-212189

Nelson, S. J., & Cha, S. (2003). Imaging glioblastoma multiforme. *Cancer J, 9*(2), 134-145. doi:10.1097/00130404-200303000-00009

Nielsen, M. (2015). Neural networks and deep learning Determination press San Francisco. *CA, USA*.

Nosouhian, S., Nosouhian, F., & Khoshouei, A. (2021). *A Review of Recurrent Neural Network Architecture for Sequence Learning: Comparison between LSTM and GRU*.

Oğuz, F. (2025). Time-based TMZ-Control Inverted Microscopic Images.

Omoseebi, A., Ali, A., Muthanna, A., Alkhalidy, M., Kumar, R. R., & Tiwari, M. (2025). Comparative Study of CNN, RNN, and Transformer Models for Sensor Data Anomaly Detection.

Osadebey, M., Liu, Q., Fuster-Garcia, E., & Emblem, K. E. (2023). Interpreting deep learning models for glioma survival classification using visualization and textual explanations. *BMC Medical Informatics and Decision Making, 23*(1), 225. doi:10.1186/s12911-023-02320-2

Ostrom, Q. T., Price, M., Neff, C., Cioffi, G., Waite, K. A., Kruchko, C., & Barnholtz-Sloan, J. S. (2023). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2016-2020. *Neuro Oncol, 25*(12 Suppl 2), iv1-iv99. doi:10.1093/neuonc/noad149

Oughtred, R., Rust, J., Chang, C., Breitkreutz, B. J., Stark, C., Willems, A., . . . Tyers, M. (2021). The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci, 30*(1), 187-200. doi:10.1002/pro.3978

Parekh, V. S., & Jacobs, M. A. (2019). Deep learning and radiomics in precision medicine. *Expert review of precision medicine and drug development, 4*(2), 59-72.

Parzen, E. (1961). An approach to time series analysis. *The Annals of Mathematical Statistics, 32*(4), 951-989.

Pati, S., Baid, U., Edwards, B., Sheller, M., Wang, S., & Reina, G. (2022). Federated learning enables big data for rare cancer boundary detection Nat. *Commun, 13*(1), 1-17.

Pavithra, M., Saruladha, K., & Sathyabama, K. (2019). *GRU based deep learning model for prognosis prediction of disease progression.* Paper presented at the 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC).

Piantino, M., Figarol, A., & Matsusaki, M. (2021). Three-Dimensional in vitro Models of Healthy and Tumor Brain Microvasculature for Drug and Toxicity Screening. *Front Toxicol, 3*, 656254. doi:10.3389/ftox.2021.656254

Prakash, K. B., & Kanagachidambaresan, G. (2021). Programming with TensorFlow. *EAI/Springer Innovations in Communication and Computing*.

Praska, C. E., Odde, D. J., & Neil, E. C. (2021). A Physical Perspective on Oncology Research: The Critically Emerging Role of Physical Science in the Fight Against Brain Cancer. *Advances in Oncology, 1*, 213-221. doi:10.1016/j.yao.2021.02.018

Pudikov, A., & Brovko, A. (2020). *Comparison of LSTM and GRU recurrent neural network architectures.* Paper presented at the International Scientific and Practical Conference in Control Engineering and Decision Making.

Qin, X., & Xu, R. (2025). Efficient Post-Quantum Cross-Silo Federated Learning Based on Key Homomorphic Pseudo-Random Function. *Mathematics, 13*(9). doi:10.3390/math13091404

Rajdeo, P., Aronow, B., & Surya Prasath, V. B. (2024). Chapter One - Deep learning-based multimodal spatial transcriptomics analysis for cancer. In E. Madan, P. B. Fisher, & R. Gogna (Eds.), *Advances in Cancer Research* (Vol. 163, pp. 1-38): Academic Press.

Ranzenberger, L. R., Das, J. M., & Snyder, T. (2025). Diffusion Tensor Imaging. In *StatPearls*. Treasure Island (FL) ineligible companies. Disclosure: Joe Das declares no relevant financial relationships with ineligible companies. Disclosure: Travis Snyder declares no relevant financial relationships with ineligible companies.: StatPearls Publishing

Ratliff, M., Karimian-Jazi, K., Hoffmann, D. C., Rauschenbach, L., Simon, M., Hai, L., . . . Winkler, F. (2023). Individual glioblastoma cells harbor both proliferative and invasive capabilities during tumor progression. *Neuro Oncol, 25*(12), 2150-2162. doi:10.1093/neuonc/noad109

Rebelo, S. P., Pinto, C., Martins, T. R., Harrer, N., Estrada, M. F., Loza-Alvarez, P., . . . Brito, C. (2018). 3D-3-culture: A tool to unveil macrophage plasticity in the tumour microenvironment. *Biomaterials, 163*, 185-197. doi:10.1016/j.biomaterials.2018.02.030

Rezaeyan, A., Asadi, S., Kamrava, S. K., & Zare-Sadeghi, A. (2023). Brain structural analysis in patients with post-traumatic anosmia: Voxel-based and surface-based morphometry. *J Neuroradiol, 50*(5), 482-491. doi:10.1016/j.neurad.2022.11.005

Rieke, N., Hancox, J., Li, W., Milletarì, F., Roth, H. R., Albarqouni, S., . . . Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digit Med, 3*, 119. doi:10.1038/s41746-020-00323-1

Rončević, A., Koruga, N., Soldo Koruga, A., & Rončević, R. (2025). Artificial Intelligence in Glioblastoma-Transforming Diagnosis and Treatment. *Chin Neurosurg J, 11*(1), 10. doi:10.1186/s41016-025-00399-2

Rončević, A., Koruga, N., Soldo Koruga, A., & Rončević, R. (2025). Artificial Intelligence in Glioblastoma, Transforming Diagnosis and Treatment. *Chinese Neurosurgical Journal, 11*(1), 10. doi:10.1186/s41016-025-00399-2

Sadegh-Zadeh, S.-A., & Bahrami, M. (2025). A review of neuroscience-inspired frameworks for machine consciousness. *Artificial Intelligence in Health*, 1. doi:10.36922/aih.5690

Samartha, M. V. S., Dubey, N. K., Jena, B., Maheswar, G., Lo, W. C., & Saxena, S. (2024). AI-driven estimation of O6 methylguanine-DNA-methyltransferase (MGMT) promoter methylation in glioblastoma patients: a systematic review with bias analysis. *J Cancer Res Clin Oncol, 150*(2), 57. doi:10.1007/s00432-023-05566-5

Sammut, S. J., Crispin-Ortuzar, M., Chin, S. F., Provenzano, E., Bardwell, H. A., Ma, W., . . . Caldas, C. (2022). Multi-omic machine learning predictor of breast cancer therapy response. *Nature, 601*(7894), 623-629. doi:10.1038/s41586-021-04278-5

Sansone, M., Fusco, R., Grassi, F., Gatta, G., Belfiore, M. P., Angelone, F., . . . Galdiero, R. (2023). Machine learning approaches with textural features to calculate breast density on mammography. *Current Oncology, 30*(1), 839-853.

Sapankevych, N. I., & Sankar, R. (2009). Time Series Prediction Using Support Vector Machines: A Survey. *IEEE Computational Intelligence Magazine, 4*(2), 24-38. doi:10.1109/MCI.2009.932254

Schapke, J., Tavares, A., & Recamonde-Mendoza, M. (2022). EPGAT: Gene Essentiality Prediction With Graph Attention Networks. *IEEE/ACM Trans Comput Biol Bioinform, 19*(3), 1615-1626. doi:10.1109/tcbb.2021.3054738

Scheins, J. J., Vahedipour, K., Pietrzyk, U., & Shah, N. J. (2015). High performance volume-of-intersection projectors for 3D-PET image reconstruction based on polar symmetries and SIMD vectorisation. *Phys Med Biol, 60*(24), 9349-9375. doi:10.1088/0031-9155/60/24/9349

Schmidgall, S., Ziaei, R., Achterberg, J., Kirsch, L., Hajiseyedrazi, S. P., & Eshraghian, J. (2024). Brain-inspired learning in artificial neural networks: A review. *APL Machine Learning, 2*(2). doi:10.1063/5.0186054

Scimeca, M., Urbano, N., Toschi, N., Bonanno, E., & Schillaci, O. (2021). Precision medicine in breast cancer: From biological imaging to artificial intelligence. *Semin Cancer Biol, 72*, 1-3. doi:10.1016/j.semcancer.2021.04.019

Serrano, D. R., Luciano, F. C., Anaya, B. J., Ongoren, B., Kara, A., Molina, G., . . . Lalatsa, A. (2024). Artificial Intelligence (AI) Applications in Drug Discovery and Drug Delivery: Revolutionizing Personalized Medicine. *Pharmaceutics, 16*(10). doi:10.3390/pharmaceutics16101328

Seyfoori, A., Liu, K., Caruncho, H. J., Walter, P. B., & Akbari, M. (2025). Tumoroid-On-a-Plate (ToP): Physiologically Relevant Cancer Model Generation and Therapeutic Screening. *Adv Healthc Mater, 14*(1), e2402060. doi:10.1002/adhm.202402060

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*: Cambridge university press.

Sharma, H. (2019). HPC-Enhanced Training of Large AI Models in the Cloud. *International Journal of Advanced Research in Engineering and Technology, 10*(2), 953-972.

Sharma, P., Aaroe, A., Liang, J., & Puduvalli, V. K. (2023). Tumor microenvironment in glioblastoma: Current and emerging concepts. *Neurooncol Adv, 5*(1), vdad009. doi:10.1093/noajnl/vdad009

Shin, H.-C., Tenenholtz, N. A., Rogers, J. K., Schwarz, C. G., Senjem, M. L., Gunter, J. L., . . . Michalski, M. (2018, 2018//). *Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks.* Paper presented at the Simulation and Synthesis in Medical Imaging, Cham.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data, 6*(1), 1-48.

Shukla, G., Alexander, G. S., Bakas, S., Nikam, R., Talekar, K., Palmer, J. D., & Shi, W. (2017). Advanced magnetic resonance imaging in glioblastoma: a review. *Chin Clin Oncol, 6*(4), 40. doi:10.21037/cco.2017.06.28

Sidey-Gibbons, J. A., & Sidey-Gibbons, C. J. (2019). Machine learning in medicine: a practical introduction. *BMC medical research methodology, 19*, 1-18.

Simon, T., Jackson, E., & Giamas, G. (2020). Breaking through the glioblastoma micro-environment via extracellular vesicles. *Oncogene, 39*(23), 4477-4490. doi:10.1038/s41388-020-1308-2

Smirnova, L. (2024). Biocomputing with organoid intelligence. *Nature Reviews Bioengineering, 2*(8), 633-634. doi:10.1038/s44222-024-00200-6

Solak, N., Ferreira, A., Luijten, G., Puladi, B., Alves, V., & Egger, J. (2025). GBM-Reservoir: Brain tumor (Glioblastoma Multiforme) MRI dataset collection with ground truth segmentation masks. *Data Brief, 58*, 111287. doi:10.1016/j.dib.2025.111287

Sonata, I., & Heryadi, Y. (2024, 17-18 July 2024). *Comparison of LSTM and Transformer for Time Series Data Forecasting.* Paper presented at the 2024 7th International Conference on Informatics and Computational Sciences (ICICoS).

Song, R., Wang, H., & Xing, Y. (2024). Retracted: Deep Learning-Based Glaucoma Detection Using CNN and Digital Fundus Images: A Promising Approach for Precise Diagnosis. *Curr Med Imaging, 20*, e15734056257657. doi:10.2174/0115734056257657231115051020

Staneva, R., Barbazan, J., Simon, A., Vignjevic, D. M., & Krndija, D. (2018). Cell Migration in Tissues: Explant Culture and Live Imaging. *Methods Mol Biol, 1749*, 163-173. doi:10.1007/978-1-4939-7701-7_13

Suresh, S. A., Vyas, S., Chu, C. H., Yamaguchi, T., Tanaka, T., Yeh, J. A., . . . Luo, Y. (2025). All-Dielectric Meta-Microlens-Array Confocal Fluorescence Microscopy. *Laser & Photonics Reviews, 19*(6), 2401314. doi:https://doi.org/10.1002/lpor.202401314

Tang, M., Jiang, S., Huang, X., Ji, C., Gu, Y., Qi, Y., . . . Yao, Y. (2024). Integration of 3D bioprinting and multi-algorithm machine learning identified glioma susceptibilities and microenvironment characteristics. *Cell Discovery, 10*(1), 39. doi:10.1038/s41421-024-00650-7

Tang, M., Rich, J. N., & Chen, S. (2021). Biomaterials and 3D Bioprinting Strategies to Model Glioblastoma and the Blood-Brain Barrier. *Adv Mater, 33*(5), e2004776. doi:10.1002/adma.202004776

Taylor-Whiteley, T. R., Le Maitre, C. L., Duce, J. A., Dalton, C. F., & Smith, D. P. (2019). Recapitulating Parkinson's disease pathology in a three-dimensional human neural cell culture model. *Dis Model Mech, 12*(4). doi:10.1242/dmm.038042

Thakuri, P. S., Liu, C., Luker, G. D., & Tavana, H. (2018). Biomaterials-Based Approaches to Tumor Spheroid and Organoid Modeling. *Advanced Healthcare Materials, 7*(6), 1700980. doi:https://doi.org/10.1002/adhm.201700980

Tolstoluzka, O., & Telezhenko, D. (2024). Development and training of LSTM models for control of virtual distributed systems using TensorFlow and Keras. *2024, 2024*(3), 11. doi:10.32620/reks.2024.3.02

Tracqui, P., & Mendjeli, M. (1999). MODELLING THREE-DIMENSIONAL GROWTH OF BRAIN TUMOURS FROM TIME SERIES OF SCANS. *Mathematical Models and Methods in Applied Sciences, 09*(04), 581-598. doi:10.1142/S0218202599000300

Tripathi, S., Gabriel, K., Dheer, S., Parajuli, A., Augustin, A. I., Elahi, A., . . . Dako, F. (2023). Understanding Biases and Disparities in Radiology AI Datasets: A Review. *J Am Coll Radiol, 20*(9), 836-841. doi:10.1016/j.jacr.2023.06.015

UCSF, B. T. C. (2024). Glioblastoma. Retrieved from https://braintumorcenter.ucsf.edu/condition/glioblastoma-grade-iv

Ullman, S. (2019). Using neuroscience to develop artificial intelligence. *science, 363*(6428), 692-693. doi:10.1126/science.aau6595

University of Texas, M. A. C. C. Glioma vs. glioblastoma: What's the difference? Retrieved from https://www.mdanderson.org/cancerwise/glioma-vs--glioblastoma--what-is-the-difference-in-these-brain-tumors-treatment-diagnosis.h00-159537378.html

Vafaei, E., & Hosseini, M. (2025). Transformers in EEG Analysis: A Review of Architectures and Applications in Motor Imagery, Seizure, and Emotion Classification. *Sensors, 25*(5), 1293. Retrieved from https://www.mdpi.com/1424-8220/25/5/1293

Van Dyk, D. A., & Meng, X.-L. (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics, 10*(1), 1-50.

Vicini, S., Bortolotto, C., Rengo, M., Ballerini, D., Bellini, D., Carbone, I., . . . Faggioni, L. (2022). A narrative review on current imaging applications of artificial intelligence and radiomics in oncology: Focus on the three most common cancers. *La radiologia medica, 127*(8), 819-836.

Walsh, L. A., & Quail, D. F. (2023). Decoding the tumor microenvironment with spatial technologies. *Nature Immunology, 24*(12), 1982-1993. doi:10.1038/s41590-023-01678-9

Wang, J., Zhang, W., Yang, H., Yeh, C. M., & Wang, L. (2022). Visual Analytics for RNN-Based Deep Reinforcement Learning. *IEEE Trans Vis Comput Graph, 28*(12), 4141-4155. doi:10.1109/tvcg.2021.3076749

Wang, L. L., Serrano, C., Zhong, X., Ma, S., Zou, Y., & Zhang, C. L. (2021). Revisiting astrocyte to neuron conversion with lineage tracing in vivo. *Cell, 184*(21), 5465-5481.e5416. doi:10.1016/j.cell.2021.09.005

Wanigasekara, J., Cullen, P. J., Bourke, P., Tiwari, B., & Curtin, J. F. (2023). Advances in 3D culture systems for therapeutic discovery and development in brain cancer. *Drug Discov Today, 28*(2), 103426. doi:10.1016/j.drudis.2022.103426

Way, G. P., Allaway, R. J., Bouley, S. J., Fadul, C. E., Sanchez, Y., & Greene, C. S. (2017). A machine learning classifier trained on cancer transcriptomes detects NF1 inactivation signal in glioblastoma. *BMC genomics, 18*, 1-11.

Weigend, A. S. (2013). Time series analysis and prediction. In *Mathematical perspectives on neural networks* (pp. 395-449): Psychology Press.

Wesseling, P., & Capper, D. (2018). WHO 2016 Classification of gliomas. *Neuropathol Appl Neurobiol, 44*(2), 139-150. doi:10.1111/nan.12432

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data, 3*, 160018. doi:10.1038/sdata.2016.18

Wu, S., Lu, L., Zhou, J., Ran, D., Wang, S., Xu, Q., . . . Lu, W. (2022). All-stage targeted therapy for glioblastoma based on lipid membrane coated cabazitaxel nanocrystals. *J Control Release, 345*, 685-695. doi:10.1016/j.jconrel.2022.03.047

Xie, T., Huang, A., Yan, H., Ju, X., Xiang, L., & Yuan, J. (2024). Artificial intelligence: illuminating the depths of the tumor microenvironment. *Journal of Translational Medicine, 22*(1), 799. doi:10.1186/s12967-024-05609-6

Xie, Z., Bailey, A., Kuleshov, M. V., Clarke, D. J. B., Evangelista, J. E., Jenkins, S. L., . . . Ma'ayan, A. (2021). Gene Set Knowledge Discovery with Enrichr. *Curr Protoc, 1*(3), e90. doi:10.1002/cpz1.90

Xu, J., Glicksberg, B. S., Su, C., Walker, P., Bian, J., & Wang, F. (2021). Federated Learning for Healthcare Informatics. *J Healthc Inform Res, 5*(1), 1-19. doi:10.1007/s41666-020-00082-4

Yan, J., Qu, W., Li, X., Wang, R., & Tan, J. (2024). GATLGEMF: A graph attention model with line graph embedding multi-complex features for ncRNA-protein interactions prediction. *Computational Biology and Chemistry, 108*, 108000. doi:https://doi.org/10.1016/j.compbiolchem.2023.108000

Yan, Y., Yang, C., Chen, W., Jia, Z., Zhou, H., Di, Z., & Xu, L. (2025). Multimodal MRI and artificial intelligence: Shaping the future of glioma. *Journal of Neurorestoratology, 13*(2), 100175. doi:https://doi.org/10.1016/j.jnrt.2024.100175

Yang, C. R., Liang, C. T., Tsai, S. C., Wu, Y. C., Liu, C. W., Yang, H. H., . . . Ma, W. L. (2022). A Novel 3D Culture Scaffold to Shorten Development Time for Multicellular Tumor Spheroids. *Int J Mol Sci, 23*(22). doi:10.3390/ijms232213962

Yang, J., Fischer, N. G., & Ye, Z. (2024). Revolutionising oral organoids with artificial intelligence. *Biomater Transl, 5*(4), 372-389. doi:10.12336/biomatertransl.2024.04.004

Yang, Y., Liu, Y., Liu, X., Gulhane, A., Mastrodicasa, D., Wu, W., . . . Patel, S. Demographic bias of expert-level vision-language foundation models in medical imaging. *Science Advances, 11*(13), eadq0305. doi:10.1126/sciadv.adq0305

Yuan, Y., & Bar-Joseph, Z. (2020). GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data. *Genome Biol, 21*(1), 300. doi:10.1186/s13059-020-02214-w

Yushchenko, A., Smelyakov, K., & Chupryna, A. (2025, 24-24 April 2025). *Evaluating CNN, RNN, and Vision Transformer for Emotion Recognition: Strengths and Weaknesses.* Paper presented at the 2025 IEEE Open Conference of Electrical, Electronic and Information Sciences (eStream).

Zaccone, G., & Karim, M. R. (2018). *Deep Learning with TensorFlow: Explore neural networks and build intelligent systems with Python*: Packt Publishing Ltd.

Zargar, S. (2021). Introduction to sequence learning models: RNN, LSTM, GRU. *Department of Mechanical and Aerospace Engineering, North Carolina State University, 37988518*.

Zhang, N., Dai, Z., Wu, W., Wang, Z., Cao, H., Zhang, Y., . . . Cheng, Q. (2021). The Predictive Value of Monocytes in Immune Microenvironment and Prognosis of Glioma Patients Based on Machine Learning. *Front Immunol, 12*, 656541. doi:10.3389/fimmu.2021.656541

Zhang, T., Zhao, X., Sun, H., Gao, B., & Liu, X. (2024). GATv2EPI: Predicting Enhancer-Promoter Interactions with a Dynamic Graph Attention Network. *Genes (Basel), 15*(12). doi:10.3390/genes15121511

Zhang, Z. (2016). A gentle introduction to artificial neural networks. *Annals of translational medicine, 4*(19).

Zhang, Z., Song, W., Wu, Q., Sun, W., Li, Q., & Jia, L. (2025). A novel local enhanced channel self-attention based on Transformer for industrial remaining useful life prediction. *Engineering Applications of Artificial Intelligence, 141*, 109815. doi:https://doi.org/10.1016/j.engappai.2024.109815

Zhao, Z., Yeoh, P. S. Q., Zuo, X., Chuah, J. H., Chow, C.-O., Wu, X., & Lai, K. W. (2024). Vision transformer-equipped Convolutional Neural Networks for automated Alzheimer's disease diagnosis using 3D MRI scans. *Frontiers in Neurology, Volume 15 - 2024*. doi:10.3389/fneur.2024.1490829

Zhu, M., Li, S., Kuang, Y., Hill, V. B., Heimberger, A. B., Zhai, L., & Zhai, S. (2022). Artificial intelligence in the radiomic analysis of glioblastomas: A review, taxonomy, and perspective. *Frontiers in Oncology, 12*, 924245.

Zou, J., & Schiebinger, L. (2021). Ensuring that biomedical AI benefits diverse populations. *EBioMedicine, 67*, 103358. doi:10.1016/j.ebiom.2021.103358

APPENDIX: SUPPLEMENTARY MATERIALS

Supplementary Material I: Python Data-Analysis and Visualization Script

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from sklearn.linear_model import LinearRegression
import os

# ================================================
# 1) READ EXCEL
# ================================================
df = pd.read_excel("RESULTS_WH.xlsx", sheet_name="MTT")

# ================================================
# 2) FIX SCALING
# ================================================
result_cols = ["RESULT1", "RESULT2", "RESULT3"]

for col in result_cols:
    df[col] = df[col].astype(float)
    df.loc[df[col] > 1000, col] = df.loc[df[col] > 1000, col] /
1000.0

df["RESULT_AVG"] = df[result_cols].mean(axis=1)

# ================================================
# 3) CREATE SAVE DIRECTORY
# ================================================
save_dir = "WH_PLOTS"
os.makedirs(save_dir, exist_ok=True)

# ================================================
# 4) FUNCTION: Plot & Predict
# ================================================
def plot_and_predict(df_sub, title, save_name):

    plt.figure(figsize=(10, 6))
    doses = sorted(df_sub["TMZ DOSE"].unique())
    cmap = plt.get_cmap("tab10")
```

```
    for i, dose in enumerate(doses):

        df_dose = df_sub[df_sub["TMZ DOSE"] ==
dose].sort_values("SAAT")

        # --- Real data ---
        plt.plot(df_dose["SAAT"], df_dose["RESULT_AVG"],
                 marker="o", markersize=8, linewidth=2,
                 color=cmap(i), label=f"{dose} mg")

        # Regression
        X = df_dose[["SAAT"]].astype(float).values
        y = df_dose["RESULT_AVG"].values

        if len(df_dose) >= 2:
            model = LinearRegression()
            model.fit(X, y)

            future_hours = np.array([[96], [122], [144]])
            preds = model.predict(future_hours)

            # --- Prediction dotted line ---
            plt.plot(future_hours.flatten(), preds,
                     linestyle='--', linewidth=2, color=cmap(i))

            # --- Prediction points as DIAMOND markers ---
            plt.scatter(future_hours.flatten(), preds,
                        marker="D", s=70, color=cmap(i))

    # Formatting
    plt.title(title)
    plt.xlabel("Saat (h)")
    plt.ylabel("RESULT_AVG")
    plt.grid(True)
    plt.legend(title="TMZ Dozu")
    plt.tight_layout()

    # SAVE PLOT
    save_path = os.path.join(save_dir, f"{save_name}.png")
    plt.savefig(save_path, dpi=300)
    plt.close()

    print(f"Saved: {save_path}")

# ================================================
# 5) SUBPLOT for 2D & 3D (same tumor type)
# ================================================
def subplot_2D_3D(df_2d, df_3d, tumor_type):
```

```
    fig, axs = plt.subplots(1, 2, figsize=(15, 6))
    cmap = plt.get_cmap("tab10")

    for ax_idx, (df_sub, title) in enumerate([(df_2d, "2D"),
(df_3d, "3D")]):

        doses = sorted(df_sub["TMZ DOSE"].unique())

        for i, dose in enumerate(doses):
            df_dose = df_sub[df_sub["TMZ DOSE"] ==
dose].sort_values("SAAT")

            # Real data
            axs[ax_idx].plot(df_dose["SAAT"],
df_dose["RESULT_AVG"],
                             marker="o", markersize=8, linewidth=2,
                             color=cmap(i), label=f"{dose} mg")

            # Regression + predictions
            X = df_dose[["SAAT"]].astype(float).values
            y = df_dose["RESULT_AVG"].values

            if len(df_dose) >= 2:
                model = LinearRegression()
                model.fit(X, y)

                future_hours = np.array([[96], [122], [144]])
                preds = model.predict(future_hours)

                axs[ax_idx].plot(future_hours.flatten(), preds,
                                 linestyle='--', linewidth=2,
color=cmap(i))

                axs[ax_idx].scatter(future_hours.flatten(), preds,
                                    marker="D", s=70,
color=cmap(i))

        axs[ax_idx].set_title(f"{tumor_type} {title}")
        axs[ax_idx].set_xlabel("Saat (h)")
        axs[ax_idx].set_ylabel("RESULT_AVG")
        axs[ax_idx].grid(True)
        axs[ax_idx].legend(title="TMZ Dozu")

    fig.tight_layout()

    # Save subplot
    save_path = os.path.join(save_dir,
f"{tumor_type}_2D_3D_subplot.png")
    fig.savefig(save_path, dpi=300)
```

```
    plt.close()

    print(f"Saved SUBPLOT: {save_path}")

# ================================================
# 6) FILTER GROUPS
# ================================================
df_A172_2D = df[(df["Tümor Tipi"] == "A172") & (df["2D/3D"] ==
"2D")]
df_A172_3D = df[(df["Tümor Tipi"] == "A172") & (df["2D/3D"] ==
"3D")]

df_U87_2D  = df[(df["Tümor Tipi"] == "U87") & (df["2D/3D"] ==
"2D")]
df_U87_3D  = df[(df["Tümor Tipi"] == "U87") & (df["2D/3D"] ==
"3D")]

# ================================================
# 7) RUN NORMAL PLOTS
# ================================================
plot_and_predict(df_A172_2D, "A172 2D", "A172_2D")
plot_and_predict(df_A172_3D, "A172 3D", "A172_3D")

plot_and_predict(df_U87_2D, "U87 2D", "U87_2D")
plot_and_predict(df_U87_3D, "U87 3D", "U87_3D")

# ================================================
# 8) RUN SUBPLOT FIGURES
# ================================================
subplot_2D_3D(df_A172_2D, df_A172_3D, "A172")
subplot_2D_3D(df_U87_2D, df_U87_3D, "U87")


# ================================================
# 9) CREATE 2x2 FINAL COMBINED SUBPLOT (A172 + U87)
# ================================================
def combine_subplots_2x2():

    # Load individual plot images
    img_A172_2D = mpimg.imread(os.path.join(save_dir,
"A172_2D.png"))
    img_A172_3D = mpimg.imread(os.path.join(save_dir,
"A172_3D.png"))
    img_U87_2D = mpimg.imread(os.path.join(save_dir, "U87_2D.png"))
    img_U87_3D = mpimg.imread(os.path.join(save_dir, "U87_3D.png"))

    # ---- Create 2x2 Figure ----
```

```
    fig, axs = plt.subplots(2, 2, figsize=(16, 12)) # Adjusted
figsize for better display

    # List of axes, images, and titles
    panels = [
        (axs[0, 0], img_A172_2D, "A172 2D"),
        (axs[0, 1], img_A172_3D, "A172 3D"),
        (axs[1, 0], img_U87_2D, "U87 2D"),
        (axs[1, 1], img_U87_3D, "U87 3D")
    ]

    for ax, img, title in panels:
        ax.imshow(img)
        ax.axis("off")
        ax.set_title(title, fontsize=14)

        # --- Add dark blue border ---
        for spine in ax.spines.values():
            spine.set_edgecolor("#001a66")
            spine.set_linewidth(1.5)
fig, axes = plt.subplots(2, 2, figsize=(12, 10))

# --- Boşluk ayarları ---
fig.subplots_adjust(
    wspace=0.05,   # sağ-sol grafikler arası mesafe
    hspace=0.08    # alt-üst grafikler arası mesafe (öncekine göre
azaltıldı)
)

# Çerçeve rengi ve kalınlığı
for ax_row in axes:
    for ax in ax_row:
        for spine in ax.spines.values():
            spine.set_edgecolor("navy")
            spine.set_linewidth(1.2)

    fig.tight_layout()

    final_path = os.path.join(save_dir,
"A172_U87_COMBINED_2x2.png")
    fig.savefig(final_path, dpi=300)
    plt.close()

    print(f"Saved FINAL 2x2 COMBINED FIGURE WITH BORDER:
{final_path}")


# Run it
combine_subplots_2x2()
```

```
# Second version v2: Negative predictions are clipped to zero

import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from sklearn.linear_model import LinearRegression

# ====================================
# 1) Read MTT sheet
# ====================================
df = pd.read_excel("RESULTS_WH.xlsx", sheet_name="MTT")

# ====================================
# 2) Fix scaling (values >1000 → /1000)
# ====================================
result_cols = ["RESULT1", "RESULT2", "RESULT3"]
for col in result_cols:
    df[col] = df[col].astype(float)
    df.loc[df[col] > 1000, col] = df.loc[df[col] > 1000, col] /
1000.0

# ====================================
# 3) Compute average
# ====================================
df["RESULT_AVG"] = df[result_cols].mean(axis=1)

# ====================================
# 4) Plot function for ONE panel
# ====================================
def plot_group(ax, df_sub, title):

    doses = sorted(df_sub["TMZ DOSE"].unique())
    cmap = plt.get_cmap("tab10")

    for i, dose in enumerate(doses):

        df_dose = df_sub[df_sub["TMZ DOSE"] ==
dose].sort_values("SAAT")

        # Measured values
        ax.plot(
            df_dose["SAAT"], df_dose["RESULT_AVG"],
            marker="o", linewidth=2, color=cmap(i),
            label=f"{dose} mg"
        )

        # Regression
        if len(df_dose) >= 2:
```

```
            X = df_dose[["SAAT"]].values
            y = df_dose["RESULT_AVG"].values

            model = LinearRegression()
            model.fit(X, y)

            future_hours = np.array([[96], [122], [144]])
            preds = model.predict(future_hours)

            #  Clip negative predictions to zero
            preds = np.maximum(preds, 0)

            ax.plot(
                future_hours.flatten(), preds,
                linestyle="--", marker="x", color=cmap(i)
            )

    ax.set_title(title)
    ax.set_xlabel("Time (h)")
    ax.set_ylabel("RESULT_AVG")
    ax.set_xlim(left=24)
    ax.set_ylim(bottom=0)
    ax.grid(True)

# ====================================
# 5) Filter groups
# ====================================
groups = [
    ("A172 2D", df[(df["Tümor Tipi"] == "A172") & (df["2D/3D"] ==
"2D")]),
    ("A172 3D", df[(df["Tümor Tipi"] == "A172") & (df["2D/3D"] ==
"3D")]),
    ("U87 2D",  df[(df["Tümor Tipi"] == "U87")  & (df["2D/3D"] ==
"2D")]),
    ("U87 3D",  df[(df["Tümor Tipi"] == "U87")  & (df["2D/3D"] ==
"3D")]),
]

# ====================================
# 6) Create combined 2×2 panel
# ====================================
fig, axes = plt.subplots(2, 2, figsize=(14, 12))
axes = axes.flatten()

for ax, (title, df_group) in zip(axes, groups):
    plot_group(ax, df_group, title)

# Shared legend
handles, labels = axes[0].get_legend_handles_labels()
```

```
fig.legend(
    handles, labels,
    title="TMZ Dose",
    loc="lower center",
    ncol=6,
    frameon=False
)

plt.tight_layout(rect=[0, 0.05, 1, 1])
plt.savefig("Combined_TimeCourse_With_Clipped_Predictions.png",
dpi=300)
plt.show()


# Third version v2: calculating Prediction of 72th hour and
accuracy

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_absolute_error,
mean_squared_error

# ====================================
# 1) READ DATA
# ====================================
df = pd.read_excel("RESULTS_WH.xlsx", sheet_name="MTT")

# ====================================
# 2) FIX SCALING (values >1000 → /1000)
# ====================================
result_cols = ["RESULT1", "RESULT2", "RESULT3"]
for col in result_cols:
    df[col] = df[col].astype(float)
    df.loc[df[col] > 1000, col] = df.loc[df[col] > 1000, col] /
1000.0

# ====================================
# 3) COMPUTE AVERAGE RESULT
# ====================================
df["RESULT_AVG"] = df[result_cols].mean(axis=1)

# ====================================
# 4) FUNCTION: TRAIN (24+48) → PREDICT 72
# ====================================
def predict_72_and_accuracy(df_sub, group_name, results_list):

    real_vals = []
```

```
    pred_vals = []

    for dose in sorted(df_sub["TMZ DOSE"].unique()):
        df_dose = df_sub[df_sub["TMZ DOSE"] == dose]

        train = df_dose[df_dose["SAAT"].isin([24, 48])]
        test  = df_dose[df_dose["SAAT"] == 72]

        if len(train) == 2 and len(test) == 1:
            X_train = train[["SAAT"]].values
            y_train = train["RESULT_AVG"].values

            model = LinearRegression()
            model.fit(X_train, y_train)

            pred_72 = model.predict(np.array([[72]]))[0]
            real_72 = test["RESULT_AVG"].values[0]

            real_vals.append(real_72)
            pred_vals.append(pred_72)

    # Accuracy metrics
    r2 = r2_score(real_vals, pred_vals)
    mae = mean_absolute_error(real_vals, pred_vals)
    rmse = np.sqrt(mean_squared_error(real_vals, pred_vals))

    results_list.append({
        "Group": group_name,
        "R2_72h": r2,
        "MAE_72h": mae,
        "RMSE_72h": rmse
    })

    # ====================================
    # PLOT: REAL vs PREDICTED (72h)
    # ====================================
    plt.figure(figsize=(6, 6))
    plt.scatter(real_vals, pred_vals)
    min_val = min(real_vals + pred_vals)
    max_val = max(real_vals + pred_vals)
    plt.plot([min_val, max_val], [min_val, max_val])
    plt.xlabel("Real 72h RESULT_AVG")
    plt.ylabel("Predicted 72h RESULT_AVG")
    plt.title(f"{group_name} - Real vs Predicted (72h)")
    plt.grid(True)
    plt.tight_layout()
    plt.savefig(f"{group_name.replace(' ',
'_')}_72h_prediction.png", dpi=300)
    plt.close()
```

```
# ====================================
# 5) FILTER GROUPS
# ====================================
groups = {
    "A172 2D": df[(df["Tümor Tipi"] == "A172") & (df["2D/3D"] ==
"2D")],
    "A172 3D": df[(df["Tümor Tipi"] == "A172") & (df["2D/3D"] ==
"3D")],
    "U87 2D":  df[(df["Tümor Tipi"] == "U87")  & (df["2D/3D"] ==
"2D")],
    "U87 3D":  df[(df["Tümor Tipi"] == "U87")  & (df["2D/3D"] ==
"3D")]
}

# ====================================
# 6) RUN ALL GROUPS
# ====================================
accuracy_results = []

for name, df_group in groups.items():
    predict_72_and_accuracy(df_group, name, accuracy_results)

# ====================================
# 7) SAVE COMBINED ACCURACY TABLE
# ====================================
accuracy_df = pd.DataFrame(accuracy_results)
accuracy_df.to_excel("Combined_72h_Prediction_Accuracy.xlsx",
index=False)

print("\n=== Combined 72h Prediction Accuracy ===")
print(accuracy_df)
```

Supplementary Material II: Python Script for Volume Calculation Based on 3D Images

```
import cv2
import numpy as np
import matplotlib.pyplot as plt


directions = ['xy1','xy2','z1','z2']
# Load the image
image_name = "u87bmecD7_2D"

x_axis = float(input('Please enter X axis in terms of
micrometer:'))
y_axis = float(input('Please enter Y axis in terms of
micrometer:'))
z_axis = float(input('Please enter Z axis in terms of
micrometre:'))

results = []
for idx,dir in enumerate(directions):

  image_path = image_name + '_' + dir + '.png'
  image = cv2.imread(image_path)

  # Convert to RGB (OpenCV loads as BGR)
  image_rgb = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)

  # Extract red regions using HSV thresholding
  hsv = cv2.cvtColor(image_rgb, cv2.COLOR_RGB2HSV)

  # Define red color range (two ranges to capture full hue wrap-
around)
  lower_red1 = np.array([0, 70, 50])
  upper_red1 = np.array([10, 255, 255])
  lower_red2 = np.array([170, 70, 50])
  upper_red2 = np.array([180, 255, 255])

  mask1 = cv2.inRange(hsv, lower_red1, upper_red1)
  mask2 = cv2.inRange(hsv, lower_red2, upper_red2)
  mask = cv2.bitwise_or(mask1, mask2)

  # Clean the mask
  mask_clean = cv2.medianBlur(mask, 5)

  # Calculate area in pixels
  red_pixels = np.sum(mask_clean > 0)
```

```
  # Calculate total pixels for normalization
  total_pixels = mask_clean.size

  # Calculate percentage area
  red_area_percent = (red_pixels / total_pixels)

  results.append(red_area_percent)


total_area = results[0] * x_axis * y_axis + results[1] * x_axis *
y_axis
total_area /= 2
total_height = results[2] * z_axis + results[3] * z_axis
total_height /= 2

print('Total area for red voxels:', total_area, 'micrometer^2')
print('Total height for red voxels:', total_height, 'micrometer')

print('Red voxel volume', total_area * total_height,
'micrometer^3')
# Show overlay visualization
overlay = image_rgb.copy()
overlay[mask_clean > 0] = [255, 255, 0]  # highlight red area in
yellow
blended = cv2.addWeighted(image_rgb, 0.6, overlay, 0.4, 0)

plt.imshow(blended)
plt.title(f"Detected red area: {red_area_percent:.2f}% of total")
plt.axis("off")
plt.show()

# Result summary (pixel-based, to be converted to mm² using axis
scale)
result = {
    "image_size_pixels": mask_clean.shape,
    "red_pixel_count": int(red_pixels),
    "total_pixel_count": int(total_pixels),
    "red_area_percent": float(red_area_percent)
}
result
```

Supplementary Material III: Python Script for Live Cell Imaging Proliferation Rate

```
import cv2
import numpy as np
import matplotlib.pyplot as plt
import glob
import os

# Load frames
image_paths = sorted(
    glob.glob(os.path.join("frames", "*.png")) +
    glob.glob(os.path.join("frames", "*.jpg")) +
    glob.glob(os.path.join("frames", "*.jpeg"))
)

dark_percentages = []
bright_percentages = []
cell_area_percentages = []

for path in image_paths:
    img = cv2.imread(path)
    H, W = img.shape[:2]

    # ------------------------------------------------------------
    # 1) REMOVE SCALE BAR (bottom-right 10% × 10%)
    # ------------------------------------------------------------
    bar_h = int(H * 0.10)
    bar_w = int(W * 0.10)
    img_clean = img.copy()
    img_clean[H-bar_h:H, W-bar_w:W] = 0

    # ------------------------------------------------------------
    # 2) MASK ONLY THE CIRCULAR WELL AREA
    # ------------------------------------------------------------
    mask = np.zeros((H, W), dtype=np.uint8)
    center = (W // 2, H // 2)
    radius = min(center) - 10
    cv2.circle(mask, center, radius, 255, -1)

    img_masked = cv2.bitwise_and(img_clean, img_clean, mask=mask)

    # ------------------------------------------------------------
    # 3) Gaussian blur (stabilizes threshold)
    # ------------------------------------------------------------
    blurred = cv2.GaussianBlur(img_masked, (9, 9), 0)

    # ------------------------------------------------------------
```

```
    # 4) Convert to HSV → use V-channel
    # ------------------------------------------------------------
    hsv = cv2.cvtColor(blurred, cv2.COLOR_BGR2HSV)
    V = hsv[:, :, 2]  # 0-255

    # ------------------------------------------------------------
    # 5) Otsu threshold
    # ------------------------------------------------------------
    _, otsu_mask = cv2.threshold(V, 0, 255, cv2.THRESH_BINARY +
cv2.THRESH_OTSU)

    # Make sure black=dark, white=bright (cells)
    black = (otsu_mask == 0).sum()
    white = (otsu_mask == 255).sum()
    total = black + white

    dark_percent = black / total * 100
    bright_percent = white / total * 100
    cell_area_percent = bright_percent   # biologically meaningful
signal

    dark_percentages.append(dark_percent)
    bright_percentages.append(bright_percent)
    cell_area_percentages.append(cell_area_percent)

    print(f"{os.path.basename(path)} -> Dark={dark_percent:.2f}%
Cells={cell_area_percent:.2f}%")

# ------------------------------------------------------------
# 6) Plot cell area percentage (most meaningful metric)
# ------------------------------------------------------------
plt.figure(figsize=(10,5))
plt.plot(cell_area_percentages, marker='o', label="Cell Area %
(white area)")
plt.plot(dark_percentages, marker='x', label="Dark % (background)")
plt.xlabel("Frame Index")
plt.ylabel("Percentage (%)")
plt.title("Cell Area & Darkness Over Time (Full Pipeline)")
plt.grid(True)
plt.legend()
plt.tight_layout()
plt.show()
```

Supplementary Material IV: Python Script for CTRL/TMZ Wound Closure Prediction

```
import pandas as pd
import numpy as np
from sklearn.linear_model import LinearRegression

# Veriyi DataFrame olarak oluştur
data = {
    "SAAT": ["24H","24H","24H","24H","48H","48H","48H","48H",
             "24H","24H","24H","24H","48H","48H","48H","48H"],
    "GRUP":
["CTRL","CTRL","CTRL","CTRL","CTRL","CTRL","CTRL","CTRL",
             "TMZ","TMZ","TMZ","TMZ","TMZ","TMZ","TMZ","TMZ"],
    "RESULT1":
[18.61,35.6,9.2,11.5,38,26.2,48.5,47.9,6.94,4.7,18.5,3.8,17.25,5.5,
32.7,6],
    "RESULT2":
[20.05,37.2,8.9,13.9,40.98,23.4,50.3,45.4,5.03,5.3,17.01,3.3,15.12,
4.8,30.09,7.9],
    "RESULT3":
[19.12,34.1,10.4,12.34,43.23,27.6,45.6,40.8,8.2,5.8,20.04,4.5,19.54
,7.5,28.5,5.1]
}

df = pd.DataFrame(data)
df["HOUR"] = df["SAAT"].str.replace("H", "").astype(int)
# Linear Regression ML Algorithm, supervised learning algorithm
# predicting hour 72 as of hour 24 and  hour 48
target_hour = 72
def predict_for_group(group_name):
  results={}
  group_df=df[df["GRUP"]==group_name]
  for result_col in ["RESULT1","RESULT2","RESULT3"]:
    x= group_df[["HOUR"]]
    y=group_df[result_col]
    model= LinearRegression()
    model.fit(x,y)
    pred=model.predict(np.array([[target_hour]]))[0]
    results[result_col]=pred
  return results
ctrl_pred= predict_for_group("CTRL")
tmz_pred= predict_for_group("TMZ")
print("CTRL PREDICTONS")
print(ctrl_pred)
print("TMZ PREDICTONS")
print(tmz_pred)
```

Supplementary Material V: Python Script for ML Models

```
!unzip photosMixed.zip -d photosMix
pip install timm torch torchvision

import torch
import torch.nn as nn
import timm

# --- 1. DeiT Backbone ---
deit = timm.create_model("deit_small_patch16_224", pretrained=True)
deit.reset_classifier(0)   # remove classifier

for param in deit.parameters():
    param.requires_grad = False  # freeze

# --- 2. Combined Model ---
class MultiModalRegressor(nn.Module):
    def __init__(self, img_backbone, img_dim=384, tab_dim=5,
out_dim=1):
        super().__init__()
        self.img_backbone = img_backbone
        self.fc = nn.Sequential(
            nn.Linear(img_dim + tab_dim, 128),
            nn.ReLU(),
            nn.Dropout(0.3),
            nn.Linear(128, out_dim)
        )

    def forward(self, img, tab):
        img_feats = self.img_backbone(img)        # [batch,
img_dim]
        combined = torch.cat([img_feats, tab], dim=1)  # concat
with tabular features
        out = self.fc(combined)
        return out


from torch.utils.data import Dataset
from torchvision import transforms
from PIL import Image
import pandas as pd
import os

transform = transforms.Compose([
    transforms.Resize((224,224)),
    transforms.ToTensor(),
    transforms.Normalize(mean=(0.485,0.456,0.406),
std=(0.229,0.224,0.225))
```

```
])

class MultiModalDataset(Dataset):
    def __init__(self, dataframe, img_dir, transform=None):
        self.df = dataframe
        self.img_dir = img_dir
        self.transform = transform

        # columns you want as tabular features
        self.tab_cols = [
            "Area_Pixels_Brut",
            "Area_Micro Meter Square Brut",
            "Area in Pixels Net",
            "Area in Micro Meter Square Net",
            "ExperimentSeries",
            "Day"
        ]

    def __len__(self):
        return len(self.df)

    def __getitem__(self, idx):
        row = self.df.iloc[idx]

        # image
        img_path = os.path.join(self.img_dir, row["Image_Name"])
        image = Image.open(img_path).convert("RGB")
        if self.transform:
            image = self.transform(image)

        # tabular features
        tabular = torch.tensor(row[self.tab_cols].values,
dtype=torch.float32)

        # target
        target = torch.tensor(row["Migration Capacity"],
dtype=torch.float32)

        return image, tabular, target


import pandas as pd

df = pd.read_csv("data.csv", delimiter=";")  # adjust delimiter if
needed
print(df.head())
print(df.columns.tolist())  # check column names
img_dir = "photosMix/photosMixed"   # folder where images are
stored
```

```
dataset = MultiModalDataset(df, img_dir, transform=transform)
from torch.utils.data import DataLoader

img_dir = "photosMix/photosMixed"   # root folder of extracted
images
dataset = MultiModalDataset(df, img_dir, transform=transform)

loader = DataLoader(dataset, batch_size=16, shuffle=True)

# Test
images, tabs, targets = next(iter(loader))
print("Image batch:", images.shape)    # [B, 3, 224, 224]
print("Tabular batch:", tabs.shape)    # [B, num_features]
print("Targets:", targets.shape)       # [B]


import os
from pathlib import Path
import pandas as pd
import numpy as np
from PIL import Image

import torch
import torch.nn as nn
from torch.utils.data import Dataset, DataLoader, random_split
from torchvision import transforms
import timm

# ------------------------------
# CONFIG
# ------------------------------
CSV_PATH = "data.csv"
IMG_ROOT = "photosMix/photosMixed"   # <== your fixed folder
BATCH_SIZE = 8
EPOCHS = 5
LR = 1e-4
VAL_FRACTION = 0.2
WORKERS = 0

TAB_COLS = [
    "Area_Pixels_Brut",
    "Area_Micro Meter Square Brut",
    "Area in Pixels Net",
    "Area in Micro Meter Square Net",
    "ExperimentSeries",
    "Day",
]
TARGET_COL = "Migration Capacity"
```

```
IMG_EXTS = {".png", ".jpg", ".jpeg", ".bmp", ".tif", ".tiff"}

# ----------------------------
# UTILITIES
# ----------------------------
def clean_numeric_series(s: pd.Series) -> pd.Series:
    """Remove spaces/commas and coerce to float."""
    s = s.astype(str).str.replace(" ", "",
regex=False).str.replace(",", "", regex=False)
    return pd.to_numeric(s, errors="coerce")


def build_image_map(img_root: str):
    image_map = {}
    for root, _, files in os.walk(img_root):
        for f in files:
            if Path(f).suffix.lower() in IMG_EXTS:
                image_map[f] = os.path.join(root, f)
    return image_map


# ----------------------------
# LOAD & CLEAN DATA
# ----------------------------
df = pd.read_csv(CSV_PATH, delimiter=";")
df.columns = df.columns.str.strip()
print("[INFO] Columns:", df.columns.tolist())

# Convert numeric cols
for c in TAB_COLS + [TARGET_COL]:
    df[c] = clean_numeric_series(df[c])

# Map images
image_map = build_image_map(IMG_ROOT)
print(f"[INFO] Found {len(image_map)} images in {IMG_ROOT}")

df["resolved_path"] = df["Image_Name"].map(image_map)
df = df.dropna(subset=TAB_COLS + [TARGET_COL] +
["resolved_path"]).reset_index(drop=True)
print("[INFO] Rows kept after cleaning:", len(df))

# ----------------------------
# SCALE TABULAR FEATURES
# ----------------------------
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
df[TAB_COLS] = scaler.fit_transform(df[TAB_COLS])

# ----------------------------
```

```
# DATASET
# ----------------------------
class MultiModalDataset(Dataset):
    def __init__(self, frame, tab_cols, target_col,
transform=None):
        self.df = frame.reset_index(drop=True)
        self.tab_cols = tab_cols
        self.target_col = target_col
        self.transform = transform

    def __len__(self):
        return len(self.df)

    def __getitem__(self, idx):
        row = self.df.iloc[idx]
        img = Image.open(row["resolved_path"]).convert("RGB")
        if self.transform:
            img = self.transform(img)
        tab =
torch.tensor(row[self.tab_cols].values.astype(np.float32))
        target = torch.tensor(float(row[self.target_col]),
dtype=torch.float32)
        return img, tab, target

transform = transforms.Compose([
    transforms.Resize((224,224)),
    transforms.ToTensor(),
    transforms.Normalize(mean=(0.485,0.456,0.406),
std=(0.229,0.224,0.225)),
])

val_len = max(1, int(len(df) * VAL_FRACTION))
train_len = len(df) - val_len
train_df, val_df = random_split(df, [train_len, val_len],
generator=torch.Generator().manual_seed(42))

train_ds = MultiModalDataset(df.iloc[train_df.indices], TAB_COLS,
TARGET_COL, transform)
val_ds   = MultiModalDataset(df.iloc[val_df.indices], TAB_COLS,
TARGET_COL, transform)

train_loader = DataLoader(train_ds, batch_size=BATCH_SIZE,
shuffle=True, num_workers=WORKERS)
val_loader   = DataLoader(val_ds, batch_size=BATCH_SIZE,
shuffle=False, num_workers=WORKERS)

print(f"[INFO] Train size: {len(train_ds)} | Val size:
{len(val_ds)}")
```

```
# ----------------------------
# MODEL
# ----------------------------
class MultiModalRegressor(nn.Module):
    def __init__(self, img_model, img_dim=384,
tab_dim=len(TAB_COLS), out_dim=1):
        super().__init__()
        self.img_backbone = img_model
        self.fc = nn.Sequential(
            nn.Linear(img_dim + tab_dim, 128),
            nn.ReLU(),
            nn.Dropout(0.3),
            nn.Linear(128, out_dim)
        )

    def forward(self, img, tab):
        img_feats = self.img_backbone(img)         # [B, img_dim]
        x = torch.cat([img_feats, tab], dim=1)
        return self.fc(x).squeeze(-1)

# Load DeiT backbone
deit = timm.create_model("deit_small_patch16_224", pretrained=True)
deit.reset_classifier(0)
for p in deit.parameters():
    p.requires_grad = False  # freeze

model = MultiModalRegressor(deit, img_dim=384,
tab_dim=len(TAB_COLS), out_dim=1)
device = "cuda" if torch.cuda.is_available() else "cpu"
model.to(device)

criterion = nn.MSELoss()
optimizer = torch.optim.Adam(model.parameters(), lr=LR)

# ----------------------------
# TRAINING LOOP
# ----------------------------
def evaluate(loader):
    model.eval()
    total, n = 0.0, 0
    with torch.no_grad():
        for imgs, tabs, ys in loader:
            imgs, tabs, ys = imgs.to(device), tabs.to(device),
ys.to(device)
            preds = model(imgs, tabs)
            loss = criterion(preds, ys)
            total += loss.item() * len(ys)
            n += len(ys)
    return total / max(1, n)
```

```
for epoch in range(1, EPOCHS+1):
    model.train()
    running, seen = 0.0, 0
    for imgs, tabs, ys in train_loader:
        imgs, tabs, ys = imgs.to(device), tabs.to(device),
ys.to(device)
        optimizer.zero_grad()
        preds = model(imgs, tabs)
        loss = criterion(preds, ys)
        loss.backward()
        optimizer.step()
        running += loss.item() * len(ys)
        seen += len(ys)
    train_loss = running / max(1, seen)
    val_loss = evaluate(val_loader)
    print(f"Epoch {epoch}/{EPOCHS} | Train Loss={train_loss:.6f} |
Val Loss={val_loss:.6f}")

import pandas as pd
import matplotlib.pyplot as plt

# Make a smaller frame with only needed days
df_days = df[df["Day"].isin([3, 5, 7])].copy()

# Group by experiment (using technique + series as ID)
group_cols = ["Culture_Technique", "ExperimentSeries"]

# Pivot so each row = experiment, columns = days
pivot = df_days.pivot_table(
    index=group_cols,
    columns="Day",
    values="Migration Capacity",
    aggfunc="first"   # in case of duplicates
)

# Rename for clarity
pivot = pivot.rename(columns={3: "day3", 5: "day5", 7: "day7"})

# Handle missing values: if one of day3/day5 is missing, take the
available one
# If both present, keep them; if one missing, replace with the
other
pivot["input_avg"] = pivot[["day3", "day5"]].mean(axis=1)

# Keep only rows with a valid Day 7
pivot = pivot.dropna(subset=["day7"]).reset_index()

print("[INFO] Rows prepared for prediction:", len(pivot))
```

```
print(pivot.head())

# --- Log transform ---
y_true_log = np.log1p(y_true)   # log(1 + y)
y_pred_log = np.log1p(y_pred)

# --- Accuracy metrics on log scale ---
r2 = r2_score(y_true_log, y_pred_log)
mse = mean_squared_error(y_true_log, y_pred_log)
rmse = np.sqrt(mse)
mae = mean_absolute_error(y_true_log, y_pred_log)

print("Log-transformed metrics:")
print(f"R² Score     : {r2:.4f}")
print(f"MSE          : {mse:.4f}")
print(f"RMSE         : {rmse:.4f}")
print(f"MAE          : {mae:.4f}")

# --- Plot normalized values ---
plt.figure(figsize=(12,6))
plt.plot(results.index, y_true_log, marker="o", color="blue",
label="True Day 7 (log)")
plt.plot(results.index, y_pred_log, marker="x", color="red",
label="Predicted Day 7 (log)")

plt.xlabel("Sample Index")
plt.ylabel("Log(Migration Capacity)")
plt.title("True vs Predicted (Log Scale)")
plt.legend()
plt.grid(True, linestyle="--", alpha=0.6)
plt.show()


# Ensure Day is integer
df["Day"] = pd.to_numeric(df["Day"],
errors="coerce").astype("Int64")

# Pivot with flat column names
pivot = df[df["Day"].isin([3,5,7])].pivot_table(
    index=["Culture_Technique","ExperimentSeries"],
    columns="Day",
    values=["Image_Name","Migration Capacity",
            "Area_Pixels_Brut","Area_Micro Meter Square Brut",
            "Area in Pixels Net","Area in Micro Meter Square Net"],
    aggfunc="first"
)

# Flatten MultiIndex columns
pivot.columns = [f"{col[0]}_{col[1]}" for col in pivot.columns]
```

```
print("Pivot columns:", pivot.columns.tolist())

# Keep only experiments with Day 7 Migration Capacity
pivot = pivot.dropna(subset=["Migration Capacity_7"]).reset_index()

# Drop rows with missing Day 5
pivot = pivot.dropna(subset=["day5"]).reset_index(drop=True)

X = pivot[["day3"]].values
y = pivot["day5"].values

model = LinearRegression()
model.fit(X, y)

# Predict Day 5 to check fit
y_pred_day5 = model.predict(X)

# Estimate Day 7 by extrapolation
growth_rate = (pivot["day5"] - pivot["day3"]) / 2  # per day growth
pivot["pred_day7"] = pivot["day5"] + growth_rate * 2

print(pivot[["Culture_Technique","ExperimentSeries","day3","day5","
pred_day7"]])


# ------------------------------
# Train a simple regression model: input = average of day3/day5 →
output = day7
# ------------------------------
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_squared_error

X = pivot[["input_avg"]].values
y = pivot["day7"].values

model_day7 = LinearRegression()
model_day7.fit(X, y)
y_pred = model_day7.predict(X)

# ------------------------------
# Compare predictions vs. true values
# ------------------------------
results = pd.DataFrame({
    "true_day7": y,
    "pred_day7": y_pred,
    "input_avg": X.flatten()
})
print(results.head())
```

```
print("R²:", r2_score(y, y_pred))
print("MSE:", mean_squared_error(y, y_pred))

# ----------------------------
# Plot
# ----------------------------
plt.figure(figsize=(6,6))
plt.scatter(y, y_pred, alpha=0.7)
plt.plot([min(y), max(y)], [min(y), max(y)], "r--")
plt.xlabel("True Day 7")
plt.ylabel("Predicted Day 7")
plt.title("Prediction vs Real (Day 7)")
plt.show()

# --- Log transform ---
y_true_log = np.log1p(y_true)   # log(1 + y)
y_pred_log = np.log1p(y_pred)

# --- Accuracy metrics on log scale ---
r2 = r2_score(y_true_log, y_pred_log)
mse = mean_squared_error(y_true_log, y_pred_log)
rmse = np.sqrt(mse)
mae = mean_absolute_error(y_true_log, y_pred_log)

print("Log-transformed metrics:")
print(f"R² Score     : {r2:.4f}")
print(f"MSE          : {mse:.4f}")
print(f"RMSE         : {rmse:.4f}")
print(f"MAE          : {mae:.4f}")

# --- Plot normalized values ---
plt.figure(figsize=(12,6))
plt.plot(results.index, y_true_log, marker="o", color="blue",
label="True Day 7 (log)")
plt.plot(results.index, y_pred_log, marker="x", color="red",
label="Predicted Day 7 (log)")

plt.xlabel("Sample Index")
plt.ylabel("Log(Migration Capacity)")
plt.title("True vs Predicted (Log Scale)")
plt.legend()
plt.grid(True, linestyle="--", alpha=0.6)
plt.show()

# ==========================
# 0) Imports & Setup
# ==========================
import os
import random
```

```
import numpy as np
import pandas as pd
from PIL import Image

import torch
import torch.nn as nn
from torch.utils.data import Dataset, DataLoader
from torchvision import transforms
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import timm

SEED = 42
random.seed(SEED); np.random.seed(SEED); torch.manual_seed(SEED)
device = torch.device("cuda" if torch.cuda.is_available() else
"cpu")
print("Device:", device)

# ==========================
# 1) Load & Clean CSV
# ==========================
csv_path   = "data.csv"  # <-- change if needed
img_dir    = "photosMix/photosMixed"  # <-- change to your images
folder
target_col = "Migration Capacity"

# Columns to use as tabular features (edit to your CSV)
tab_cols = [
    "Area_Pixels_Brut",
    "Area_Micro Meter Square Brut",
    "Area in Pixels Net",
    "Area in Micro Meter Square Net",
    "ExperimentSeries",
    "Day",
]

# Read CSV
df = pd.read_csv(csv_path, delimiter=";")

# Replace decimal commas with dots across all string cells
df = df.applymap(lambda x: str(x).replace(",", ".") if
isinstance(x, str) else x)

# Coerce numerics (target + features) and drop bad rows
df[target_col] = pd.to_numeric(df[target_col], errors="coerce")
for c in tab_cols:
    df[c] = pd.to_numeric(df[c], errors="coerce")

# Keep rows with valid target & features
```

```
df = df.dropna(subset=[target_col] + tab_cols)

# Basic check
print("Cleaned shape:", df.shape)
print("Sample columns:", df.columns.tolist()[:12])

# =========================
# 2) Train/Val Split (no leakage)
# =========================
train_df, val_df = train_test_split(df, test_size=0.2,
random_state=SEED, shuffle=True)

# Scale tabular features (fit on train ONLY)
scaler = StandardScaler()
train_df[tab_cols] = scaler.fit_transform(train_df[tab_cols])
val_df[tab_cols]   = scaler.transform(val_df[tab_cols])

# (Optional) z-score target for stability (fit on train only)
y_mean = train_df[target_col].mean()
y_std  = train_df[target_col].std() if train_df[target_col].std() >
0 else 1.0
train_df[target_col] = (train_df[target_col] - y_mean) / y_std
val_df[target_col]   = (val_df[target_col] - y_mean) / y_std

# =========================
# 3) Transforms
# =========================
train_tfms = transforms.Compose([
    transforms.Resize((224, 224)),
    transforms.RandomHorizontalFlip(p=0.5),
    transforms.ToTensor(),
    transforms.Normalize(mean=(0.485,0.456,0.406),
std=(0.229,0.224,0.225)),
])
val_tfms = transforms.Compose([
    transforms.Resize((224, 224)),
    transforms.ToTensor(),
    transforms.Normalize(mean=(0.485,0.456,0.406),
std=(0.229,0.224,0.225)),
])

# =========================
# 4) Dataset
# =========================
class MultiModalDataset(Dataset):
    def __init__(self, dataframe, img_dir, tab_cols, target_col,
transform=None):
        self.df = dataframe.reset_index(drop=True)
        self.img_dir = img_dir
```

```
        self.tab_cols = tab_cols
        self.target_col = target_col
        self.transform = transform

    def __len__(self):
        return len(self.df)

    def __getitem__(self, idx):
        row  = self.df.iloc[idx]
        path = os.path.join(self.img_dir, row["Image_Name"])
        img  = Image.open(path).convert("RGB")

        if self.transform:
            img = self.transform(img)

        tabs   = torch.tensor(row[self.tab_cols].values,
dtype=torch.float32)
        target = torch.tensor(row[self.target_col],
dtype=torch.float32)
        return img, tabs, target

train_ds = MultiModalDataset(train_df, img_dir, tab_cols,
target_col, transform=train_tfms)
val_ds   = MultiModalDataset(val_df,   img_dir, tab_cols,
target_col, transform=val_tfms)

train_loader = DataLoader(train_ds, batch_size=16, shuffle=True,
num_workers=2, pin_memory=True)
val_loader   = DataLoader(val_ds,   batch_size=32, shuffle=False,
num_workers=2, pin_memory=True)

# Sanity check
imgs, tabs, tgts = next(iter(train_loader))
print("Image batch:", imgs.shape)    # [B, 3, 224, 224]
print("Tabular:", tabs.shape)        # [B, len(tab_cols)]
print("Targets:", tgts.shape)        # [B]

# =========================
# 5) Model: CNN (ResNet50) + Transformer (ViT-S) + Tabular
# =========================
class CNNTransformerRegressor(nn.Module):
    def __init__(self, tab_dim, out_dim=1):
        super().__init__()
        # CNN backbone (global pooled features)
        # timm: num_classes=0 => returns pooled feature vector
directly
        self.cnn = timm.create_model("resnet50", pretrained=True,
num_classes=0)  # 2048-dim
        cnn_dim = 2048
```

```
        # Transformer backbone (ViT small)
        self.vit = timm.create_model("vit_small_patch16_224",
pretrained=True)
        self.vit.reset_classifier(0)  # 384-dim features
        vit_dim = 384

        # (Optional) freeze encoders for stability/low data regimes
        for p in self.cnn.parameters():
            p.requires_grad = False
        for p in self.vit.parameters():
            p.requires_grad = False

        fused_dim = cnn_dim + vit_dim + tab_dim
        self.head = nn.Sequential(
            nn.Linear(fused_dim, 256),
            nn.ReLU(inplace=True),
            nn.Dropout(0.3),
            nn.Linear(256, 64),
            nn.ReLU(inplace=True),
            nn.Dropout(0.2),
            nn.Linear(64, out_dim),
        )

    def forward(self, x_img, x_tab):
        with torch.no_grad():
            cnn_feats = self.cnn(x_img)      # [B, 2048]
            vit_feats = self.vit(x_img)      # [B, 384]

        fused = torch.cat([cnn_feats, vit_feats, x_tab], dim=1)
        out = self.head(fused).squeeze(-1)  # [B] for regression
        return out

model = CNNTransformerRegressor(tab_dim=len(tab_cols),
out_dim=1).to(device)

# =========================
# 6) Loss, Optim, (optional) Scheduler
# =========================
criterion = nn.MSELoss()
optimizer = torch.optim.AdamW(model.head.parameters(), lr=1e-4,
weight_decay=1e-4)
# Only optimizing the head since backbones are frozen. If you
unfreeze later, pass model.parameters().

# =========================
# 7) Train/Eval Loops
# =========================
```

```
def train_one_epoch(model, loader, optimizer, criterion,
clip_grad=1.0):
    model.train()
    running = 0.0
    for imgs, tabs, tgts in loader:
        imgs = imgs.to(device, non_blocking=True)
        tabs = tabs.to(device, non_blocking=True)
        tgts = tgts.to(device, non_blocking=True)

        optimizer.zero_grad(set_to_none=True)
        preds = model(imgs, tabs)
        loss  = criterion(preds, tgts)
        if torch.isnan(loss) or torch.isinf(loss):
            print(" NaN/Inf loss detected; skipping this batch.")
            continue

        loss.backward()
        nn.utils.clip_grad_norm_(model.parameters(), clip_grad)
        optimizer.step()

        running += loss.item() * imgs.size(0)
    return running / len(loader.dataset)

@torch.no_grad()
def evaluate(model, loader, criterion):
    model.eval()
    running = 0.0
    for imgs, tabs, tgts in loader:
        imgs = imgs.to(device, non_blocking=True)
        tabs = tabs.to(device, non_blocking=True)
        tgts = tgts.to(device, non_blocking=True)

        preds = model(imgs, tabs)
        loss  = criterion(preds, tgts)
        running += loss.item() * imgs.size(0)
    return running / len(loader.dataset)

# =========================
# 8) Training
# =========================
EPOCHS = 10
best_val = float("inf")
for epoch in range(1, EPOCHS+1):
    tr_loss = train_one_epoch(model, train_loader, optimizer,
criterion, clip_grad=1.0)
    val_loss = evaluate(model, val_loader, criterion)
    print(f"Epoch {epoch:02d} | train MSE: {tr_loss:.4f} | val MSE:
{val_loss:.4f}")
```

```
    if val_loss < best_val:
        best_val = val_loss
        torch.save({
            "model": model.state_dict(),
            "scaler": scaler,  # for tabular inverse-transform
later if needed
            "y_mean": y_mean,
            "y_std": y_std,
            "tab_cols": tab_cols,
            "target_col": target_col,
        }, "best_cnn_vit_tabular.pt")
        print(" Saved new best model.")

print("Training done. Best val MSE:", best_val)

# =========================
# 9) Inference helper (example)
# =========================
@torch.no_grad()
def predict_batch(model, imgs, tabs):
    model.eval()
    imgs = imgs.to(device)
    tabs = tabs.to(device)
    preds = model(imgs, tabs)
    # De-normalize target back to original scale
    preds = preds * y_std + y_mean
    return preds.cpu().numpy()

#===== CNN&LSTM MODEL

class CovidPredictor(nn.Module):
    def __init__(self, n_features, n_hidden, seq_len, n_layers):
        super(CovidPredictor, self).__init__()
        self.n_hidden = n_hidden
        self.seq_len = seq_len
        self.n_layers = n_layers
        self.c1 = nn.Conv1d(in_channels=1, out_channels=1,
kernel_size = 2, stride = 1) # Add a 1D CNN layer
        self.lstm = nn.LSTM(
            input_size=n_features,
            hidden_size=n_hidden,
            num_layers=n_layers
        )
        self.linear = nn.Linear(in_features=n_hidden,
out_features=1)
    def reset_hidden_state(self):
        self.hidden = (
            torch.zeros(self.n_layers, self.seq_len-1,
self.n_hidden),
```

```
            torch.zeros(self.n_layers, self.seq_len-1,
self.n_hidden)
        )
    def forward(self, sequences):
        sequences = self.c1(sequences.view(len(sequences), 1, -1))
        lstm_out, self.hidden = self.lstm(
            sequences.view(len(sequences), self.seq_len-1, -1),
            self.hidden
        )
        last_time_step = lstm_out.view(self.seq_len-1,
len(sequences), self.n_hidden)[-1]
        y_pred = self.linear(last_time_step)
        return y_pred
--training
def train_model(model, train_data, train_labels, val_data=None,
val_labels=None, num_epochs=100, verbose = 10, patience = 10):
    loss_fn = torch.nn.L1Loss() #
    optimiser = torch.optim.Adam(model.parameters(), lr=0.001)
    train_hist = []
    val_hist = []
    for t in range(num_epochs):

        epoch_loss = 0

        for idx, seq in enumerate(train_data): # hidden state needs
to be reset after every sample

            model.reset_hidden_state()

            # train loss
            seq = torch.unsqueeze(seq, 0)
            y_pred = model(seq)
            loss = loss_fn(y_pred[0].float(), train_labels[idx]) #
calculated loss after 1 step

            # update weights
            optimiser.zero_grad()
            loss.backward()
            optimiser.step()

            epoch_loss += loss.item()

        train_hist.append(epoch_loss / len(train_data))

        if val_data is not None:

            with torch.no_grad():

                val_loss = 0
```

```
                for val_idx, val_seq in enumerate(val_data):

                    model.reset_hidden_state() # hidden state reset
every sequence
                    val_seq = torch.unsqueeze(val_seq, 0)
                    y_val_pred = model(val_seq)
                    val_step_loss = loss_fn(y_val_pred[0].float(),
val_labels[val_idx])

                    val_loss += val_step_loss

            val_hist.append(val_loss / len(val_data)) # append in
val hist

            ## print loss for every `verbose` times
            if t % verbose == 0:
                print(f'Epoch {t} train loss: {epoch_loss /
len(train_data)} val loss: {val_loss / len(val_data)}')

            ## check early stopping for every `patience` times
            if (t % patience == 0) & (t != 0):

                ## if loss increased, perform early stopping
                if val_hist[t - patience] < val_hist[t] :

                    print('\n Early Stopping')

                    break

        elif t % verbose == 0:
            print(f'Epoch {t} train loss: {epoch_loss /
len(train_data)}')

        return model, train_hist, val_hist

# CNN & HYBRID MODELS
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten,
Dense, Dropout
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.layers import TimeDistributed, LSTM, Input
from tensorflow.keras.models import Model
from PIL import Image
import os
import pandas as pd
import numpy as np

from sklearn.model_selection import train_test_split
```

```
from sklearn.preprocessing import MinMaxScaler
# Specify the folder containing .tif images
# Search for all subfolders and files that are .tif images
folder_path = 'photos'
tif_images = {} # Dictionary to store images keys are as filenames
filenames = []
for root, dirs, files in os.walk(folder_path):
    for filename in files:
        if filename.endswith('.tif'):
            image_path = os.path.join(root, filename)
            image = Image.open(image_path)
            filenames.append(filename)
            tif_images[filename] = image

print(f"Loaded {len(tif_images)} .tif images from all subfolders.")

data = pd.read_csv('data.csv',delimiter=';')
data_photos = data[data['Image_Name'].isin(filenames)]
data_photos_direct = data_photos[data_photos['Culture_Technique ']
== 'direct']
data_photos_direct.info()
data_photos_direct['Migration Capacity '] =
(data_photos_direct['Migration Capacity ']
    .str.replace(',', '.')  # Replace comma with period
    .astype(float))
migration_capacity_dict =
dict(zip(data_photos_direct['Image_Name'],
data_photos_direct['Migration Capacity ']))

data_d3_e1 =
list(data_photos_direct['Image_Name'][(data_photos_direct['Experime
ntSeries'] == 1) & (data_photos_direct['Day'] == 3)])
data_d3_e2 =
list(data_photos_direct['Image_Name'][(data_photos_direct['Experime
ntSeries'] == 2) & (data_photos_direct['Day'] == 3)])
data_d5_e1 =
list(data_photos_direct['Image_Name'][(data_photos_direct['Experime
ntSeries'] == 1) & (data_photos_direct['Day'] == 5)])
data_d5_e2 =
list(data_photos_direct['Image_Name'][(data_photos_direct['Experime
ntSeries'] == 2) & (data_photos_direct['Day'] == 5)])
data_d7_e1 =
list(data_photos_direct['Image_Name'][(data_photos_direct['Experime
ntSeries'] == 1) & (data_photos_direct['Day'] == 7)])
data_d7_e2 =
list(data_photos_direct['Image_Name'][(data_photos_direct['Experime
ntSeries'] == 2) & (data_photos_direct['Day'] == 7)])

print("Length of data_d3_e1:", len(data_d3_e1))
```

```
print("Length of data_d3_e2:", len(data_d3_e2))
print("Length of data_d5_e1:", len(data_d5_e1))
print("Length of data_d5_e2:", len(data_d5_e2))
print("Length of data_d7_e1:", len(data_d7_e1))
print("Length of data_d7_e2:", len(data_d7_e2))

sequences_e1 = []
for i1 in data_d3_e1:
    for i2 in data_d5_e1:
        for i3 in data_d7_e1:
            sequences_e1.append([i1, i2, i3])

sequences_e2 = []
for i1 in data_d3_e2:
    for i2 in data_d5_e2:
        for i3 in data_d7_e2:
            sequences_e2.append([i1, i2, i3])

sequences = sequences_e1 + sequences_e2

labels = [seq[-1] for seq in sequences]
sequences = [seq[:-1] for seq in sequences]

data_train, data_test, label_train, label_test =
train_test_split(sequences, labels, test_size=0.15)

data_train = np.array(data_train)
data_test = np.array(data_test)
label_train = np.array(label_train)
label_test = np.array(label_test)
data_train.shape , data_test.shape

def reshape_images(image_dict, target_size=(224, 224)):

reshaped_images = {}
    for filename, image in image_dict.items():
        reshaped_image = image.resize(target_size)
        reshaped_images[filename] = np.array(reshaped_image)
    return reshaped_images
photos_array = reshape_images(tif_images, target_size=(224,224))

data_train
X_train = [[photos_array[f] for f in seq] for seq in data_train]
X_train = np.array(X_train)
X_test = [[photos_array[f] for f in seq] for seq in data_test]
X_test = np.array(X_test)

y_train = np.array([migration_capacity_dict[f] for f in
label_train])
```

```
y_test = np.array([migration_capacity_dict[f] for f in label_test])

X_train[0].shape
scaler = MinMaxScaler()
y_train_scaled = scaler.fit_transform(y_train.reshape(-1, 1))
y_test_scaled = scaler.transform(y_test.reshape(-1, 1))

cnn_model = Sequential([
Conv2D(32, (3, 3), activation='relu', input_shape=(224,224, 3)),
MaxPooling2D((2, 2)),
Flatten(),
Dense(16, activation='relu')
])
input_layer = Input(shape=(2, 224,224, 3))
encoded_sequence = TimeDistributed(cnn_model)(input_layer)
# Add RNN layers
lstm_out = LSTM(32, return_sequences=False)(encoded_sequence)

# Add Dense layers for final prediction
x = Dense(16, activation='relu')(lstm_out)
output = Dense(1)(x)  # Single real value output

model = Model(inputs=input_layer, outputs=output)
model.compile(optimizer='adam', loss='mean_squared_error',
metrics=['mae'])

history = model.fit(
    X_train, y_train,
    epochs=30,
    batch_size=32,
)
predictions = model.predict(X_test)
real_predictions = scaler.inverse_transform(predictions)

mape = np.mean(np.abs((real_predictions - y_test) / y_test))
mape * 100
predictions_train = model.predict(X_train)
real_predictions_train =
scaler.inverse_transform(predictions_train)
mape_train = np.mean(np.abs((real_predictions_train - y_train) /
y_train))
mape_train * 100

import matplotlib.pyplot as plt
# Plot real values vs predictions
plt.figure(figsize=(10, 6))
plt.plot(y_test, label='Real Values', marker='o')
plt.plot(real_predictions, label='Predictions', marker='x')
plt.title('Test Predictions vs Real Values')
```

```
plt.xlabel('Sample Index')
plt.ylabel('migration capacity')
plt.legend()
plt.grid(True)
plt.show()
predictions
X_test[0]

# GRU Model
#import tensorflow as tf
print("started")
import numpy as np
import pandas as pd
print("numpy and pandas imported")

from tensorflow import keras
print("tensor imported")

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler
data = pd.read_csv('data.csv', delimiter=';')
data.info()
data = data.dropna(subset=['Day'])
data.info()
data_extracted = data[['ExperimentSeries', 'Area_Pixels_Brut',
'Day','Culture_Technique ']]

data_direct = data_extracted[data_extracted['Culture_Technique ']
== 'direct']

data_direct.info()
scaler = MinMaxScaler()
data_direct['Area_Pixels_Brut'] =
scaler.fit_transform(data_direct[['Area_Pixels_Brut']])

data_d3_e1 =
list(data_direct['Area_Pixels_Brut'][(data_direct['ExperimentSeries
'] == 1) & (data_direct['Day'] == 3)])
data_d3_e2 =
list(data_direct['Area_Pixels_Brut'][(data_direct['ExperimentSeries
'] == 2) & (data_direct['Day'] == 3)])
data_d5_e1 =
list(data_direct['Area_Pixels_Brut'][(data_direct['ExperimentSeries
'] == 1) & (data_direct['Day'] == 5)])
data_d5_e2 =
list(data_direct['Area_Pixels_Brut'][(data_direct['ExperimentSeries
'] == 2) & (data_direct['Day'] == 5)])
data_d7_e1 =
list(data_direct['Area_Pixels_Brut'][(data_direct['ExperimentSeries
'] == 1) & (data_direct['Day'] == 7)])
data_d7_e2 =
list(data_direct['Area_Pixels_Brut'][(data_direct['ExperimentSeries
'] == 2) & (data_direct['Day'] == 7)])

sequences_e1 = []
for i1 in data_d3_e1:
    for i2 in data_d5_e1:
```

```
        for i3 in data_d7_e1:
            sequences_e1.append([i1, i2, i3])

sequences_e2 = []
for i1 in data_d3_e2:
    for i2 in data_d5_e2:
        for i3 in data_d7_e2:
            sequences_e2.append([i1, i2, i3])

sequences = sequences_e1 + sequences_e2

labels = [seq[-1] for seq in sequences]
sequences = [seq[:-1] for seq in sequences]

data_train, data_test, label_train, label_test =
train_test_split(sequences, labels, test_size=0.15)

data_train = np.array(data_train)
data_test = np.array(data_test)
label_train = np.array(label_train)
label_test = np.array(label_test)
data_train.shape , data_test.shape

# Define the model
model = keras.Sequential()
model.add(keras.layers.GRU(128, activation='tanh',
return_sequences=True, input_shape=(2, 1))),
#model.add(keras.layers.Dense(64)),
model.add(keras.layers.GRU(128, activation='tanh')),
model.add(keras.layers.Dense(32))
model.add(keras.layers.Dense(1))

# Compile the model
model.compile(optimizer='adam', loss='mse')
# Print the model summary
model.summary()
sample_input = data_train[0:1]  # Take first sample to get the
shape
model(sample_input)  # This builds the model

history = model.fit(data_train, label_train, batch_size=32,
epochs=30, verbose=1)
predictions = model.predict(data_test)
real_predictions = scaler.inverse_transform(predictions)
print("predictions: ",real_predictions)
real_label_test = scaler.inverse_transform(label_test.reshape(-1,
1))
print("real_label_test: ",real_label_test)

maep = np.mean(np.abs((real_label_test - real_predictions) /
real_label_test)) * 100
print("Mean Absolute Percentage Error: ",maep)
```

Supplementary Material IV: MATLAB Image Processing Script

```
direct1_astrocytes_day7_masked = imread("/MATLAB
Drive/Astrocytes/direct1-
astrocytes_day7_masked.tif");
direct1_Astrocytes_day5 = imread("/MATLAB Drive/Astrocytes/direct1-
Astrocytes_day5.tif");
direct1_Astrocytes_day3 = imread("/MATLAB
Drive/Astrocytes/direct1_Astrocytes_day3.tif”);
imshow(direct1_Astrocytes_day3)imshow(direct1_Astrocytes_day3)
title("Astrocytes direct day 3”)
# converting images into grayscale
grayAstD3=im2gray(direct1_Astrocytes_day3);
imshow(grayAstD3)
title("Astrocytes Day 3 in gray scale”)
adjAstD3=imadjust(grayAstD3)
imhist(adjAstD3)
title("Adjusted Histogram")
imhist(adjAstD3)
#adding a circle onto an image that covers the region of the
doughnut
adjAstD3Circle = insertShape(adjAstD3,"filledcircle",[150 280
35],LineWidth=5);
imshow(adjAstD3Circle)
adjAstD3Circle = insertShape(adjAstD3,"filledcircle",[512 512
35],LineWidth=5);
imshow(adjAstD3Circle)
gray_direct1_Ast=im2gray(direct1_Astrocytes)
adj_gray_direct1_Ast=imadjust(gray_direct1_Ast)
circle_adj_gray_direct1_Ast =
insertShape(adj_gray_direct1_Ast,"filledcircle",[150
280 35],LineWidth=5)
circle_adj_gray_direct1_Ast =
insertShape(adj_gray_direct1_Ast,"filledcircle",[530
550 400],LineWidth=5)
masked_adj_gray_direct1_Ast =
insertShape(adj_gray_direct1_Ast,'filledcircle',[530
550 400],'LineWidth',5 , 'color',[1 1 1])
N=numel(circle_adj_gray_direct1_Ast)
props = regionprops("table",BW,"all”)
sum(props.Area)

bw_adj_gray_direct1_Ast = createBinaryMask(adj_gray_direct1_Ast
>> AreaPixels= sum(props_bw_adj_gray_circle_Stamping1_U87_1D5.Area)
% Adjust data to span data range.
X = imadjust(X);
% Threshold image with manual threshold
BW = im2gray(X) > 75;
% Open mask with square
width = 4;
se = strel('square', width);
BW = imopen(BW, se);
% Create masked image.
maskedImage = X;
maskedImage(~BW) = 0;
end
```